

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: July 13, 2004, 12:34:44 ; Search time 88 Seconds  
 (without alignments)  
 12108.099 Million cell updates/sec

Title: US-09-270-437D-5  
 Perfect score: 560  
 Sequence: 1 aggaagctgcgcacgcgc.....attctctcagggttttaaaa 1708

## Scoring table:

OLIGO  
 Xgapop 60.0 , Xgapext 60.0  
 Ygapop 60.0 , Ygapext 60.0  
 Fgapop 6.0 , Fgapext 7.0  
 Delop 6.0 , Delext 7.0

Searched: 1279676 seqs, 311918243 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2457832

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp  
 -Q=/cgn2\_1/USPTO\_spool\_p/US09270437/runat\_13072004\_10114/app\_query.fasta\_1.1863  
 -DB=Published Applications AA -QFT=fastan -SUFFIX=rapb -MINMATCH=0.1  
 -LOOPCL=0 -LOOPEXT=0 -UNIT5=5 -START=1 -END=1 -MATRIX=coligo  
 -TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=quality -THR\_MIN=1  
 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
 -MAXLEN=2000000000 -USER=US09270437 @CGN 1 1 130 @runat 13072004 121958 10114  
 -NCPU=6 -ICPU=3 -NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
 -LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=10 -THREADS=1 -XGAPOP=60 -XGAPEXT=60  
 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Published Applications AA:\*\*  
 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
 2: /cgn2\_6/ptodata/1/pubpaa/pct\_NEW\_PUB.pep.\*  
 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
 12: /cgn2\_6/ptodata/1/pubpaa/US09A\_NEW\_PUB.pep.\*  
 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match Length | ID | Description |
|------------|-------|--------------------|----|-------------|
| -----      |       |                    |    |             |

| 1  | 261 | 46.6 | 577 | 15 | US-10-313-986-500   | Sequence 500, App  |
|----|-----|------|-----|----|---------------------|--------------------|
| 2  | 216 | 38.6 | 261 | 9  | US-09-764-864-1114  | Sequence 1114, App |
| 3  | 166 | 29.6 | 577 | 9  | US-09-873-637-2     | Sequence 2, Appl1  |
| 4  | 116 | 20.7 | 250 | 9  | US-09-764-864-1532  | Sequence 1532, App |
| 5  | 49  | 8.8  | 49  | 9  | US-09-873-637-18    | Sequence 18, Appl  |
| 6  | 48  | 8.6  | 48  | 9  | US-09-873-637-20    | Sequence 20, Appl  |
| 7  | 47  | 8.4  | 47  | 9  | US-09-873-637-17    | Sequence 17, Appl  |
| 8  | 45  | 8.0  | 47  | 9  | US-09-873-637-19    | Sequence 19, Appl  |
| 9  | 36  | 6.4  | 70  | 14 | US-10-117-982-472   | Sequence 472, App  |
| 10 | 36  | 6.4  | 70  | 15 | US-10-313-986-472   | Sequence 472, App  |
| 11 | 36  | 6.4  | 422 | 16 | US-10-408-765A-2088 | Sequence 2088, App |
| 12 | 36  | 6.4  | 579 | 9  | US-09-735-705-176   | Sequence 176, App  |
| 13 | 36  | 6.4  | 579 | 9  | US-09-735-705-348   | Sequence 348, App  |
| 14 | 36  | 6.4  | 579 | 9  | US-09-850-716A-176  | Sequence 176, App  |
| 15 | 36  | 6.4  | 579 | 9  | US-09-850-716A-348  | Sequence 348, App  |
| 16 | 36  | 6.4  | 579 | 9  | US-09-897-778-176   | Sequence 176, App  |
| 17 | 36  | 6.4  | 579 | 9  | US-09-897-778-348   | Sequence 348, App  |
| 18 | 36  | 6.4  | 579 | 9  | US-09-897-778-446   | Sequence 446, App  |
| 19 | 36  | 6.4  | 579 | 9  | US-09-897-778-449   | Sequence 449, App  |
| 20 | 36  | 6.4  | 579 | 10 | US-09-466-396A-176  | Sequence 176, App  |
| 21 | 36  | 6.4  | 579 | 12 | US-10-007-700-176   | Sequence 176, App  |
| 22 | 36  | 6.4  | 579 | 12 | US-10-007-700-348   | Sequence 348, App  |
| 23 | 36  | 6.4  | 579 | 12 | US-10-007-700-446   | Sequence 446, App  |
| 24 | 36  | 6.4  | 579 | 12 | US-10-007-700-449   | Sequence 449, App  |
| 25 | 36  | 6.4  | 579 | 14 | US-10-117-982-176   | Sequence 176, App  |
| 26 | 36  | 6.4  | 579 | 14 | US-10-117-982-348   | Sequence 348, App  |
| 27 | 36  | 6.4  | 579 | 14 | US-10-117-982-446   | Sequence 446, App  |
| 28 | 36  | 6.4  | 579 | 14 | US-10-117-982-449   | Sequence 449, App  |
| 29 | 36  | 6.4  | 579 | 14 | US-10-117-982-480   | Sequence 480, App  |
| 30 | 36  | 6.4  | 579 | 14 | US-10-117-982-484   | Sequence 484, App  |
| 31 | 36  | 6.4  | 579 | 15 | US-10-313-986-176   | Sequence 176, App  |
| 32 | 36  | 6.4  | 579 | 15 | US-10-313-986-348   | Sequence 348, App  |
| 33 | 36  | 6.4  | 579 | 15 | US-10-313-986-446   | Sequence 446, App  |
| 34 | 36  | 6.4  | 579 | 15 | US-10-313-986-449   | Sequence 449, App  |
| 35 | 36  | 6.4  | 579 | 15 | US-10-313-986-480   | Sequence 480, App  |
| 36 | 36  | 6.4  | 579 | 15 | US-10-313-986-484   | Sequence 484, App  |
| 37 | 36  | 6.4  | 586 | 9  | US-09-850-716A-427  | Sequence 427, App  |
| 38 | 36  | 6.4  | 586 | 9  | US-09-897-778-427   | Sequence 427, App  |
| 39 | 36  | 6.4  | 586 | 12 | US-10-007-700-427   | Sequence 427, App  |
| 40 | 36  | 6.4  | 586 | 14 | US-10-117-982-427   | Sequence 427, App  |
| 41 | 36  | 6.4  | 586 | 15 | US-10-313-986-427   | Sequence 427, App  |
| 42 | 36  | 6.4  | 589 | 15 | US-10-313-986-486   | Sequence 486, App  |
| 43 | 28  | 5.0  | 47  | 9  | US-09-873-637-21    | Sequence 21, Appl  |
| 44 | 24  | 4.3  | 70  | 14 | US-10-117-982-473   | Sequence 473, App  |
| 45 | 24  | 4.3  | 70  | 15 | US-10-313-986-473   | Sequence 473, App  |

## ALIGNMENTS

RESULT 1  
 US-10-313-986-500  
 ; Sequence 500, Application US/10313986  
 ; Publication No. US20030236209A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Foy, Teresa M.  
 ; APPLICANT: McNabb, Andria  
 ; APPLICANT: Watanabe, Yoshihiro  
 ; APPLICANT: Reed, Steven G.  
 ; APPLICANT: Wang, Tongtong  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
 ; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
 ; FILE REFERENCE: 210121.455C19  
 ; CURRENT APPLICATION NUMBER: US/10/313,986  
 ; CURRENT FILING DATE: 2002-12-04  
 ; NUMBER OF SEQ ID NOS: 560  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 500  
 ; LENGTH: 577  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-313-986-500

Alignment Scores:

```
Pred. No.: 8.27e-241 Length: 577
Score: 261.00 Matches: 441
Percent Similarity: 99.55% Conservative: 0
Best Local Similarity: 99.55% Mismatches: 0
Query Match: 46.61% Indels: 2
DB: 15 Gaps: 1

US-09-270-437D-5 (1-1708) x US-10-313-986-500 (1-577)

QY 38 GCATCATGAGCTGAATGCCACACAGCTTGAGAGAACCATGCCCTGAGAGTCTCTACATC 97
Db 135 AlalileMetLysLeuAsnGlyHisGlnLeuGluAsnHisAlaLeuLysValSerTyrIle 154
QY 98 CCCGATGAGCAGATAGACACAGGACCTGAGAAATGGCGCGCCGAGGGGCTTTGGCTCTCGG 157
Db 155 ProaspGluGlnIleAlaGlnGlyProGluAsnGlyArgGlyGlyPheGlySerArg 174
QY 158 GGTCAAGCCCCCGCAGGGCTCACCTGTGCGACGGGGGCCCGCCAGCAAGCAGCAAGTG 217
Db 175 GlyGlnProArgGlnGlySerProValAlaAlaGlyAlaProAlaLysGlnGlnVal 194
QY 218 GACATCCCTCCCTCGCTCCTGCTGCCACCCAGTATGCTGGTGCATTATTGCAAGGAG 277
Db 195 AspileProLeuArgLeuValProThrGlnIleValGlyAlaIleIleGlyLysGlu 214
QY 278 GGGGCCACCATCCGCAACATCACAAACAGACCCAGTCCAAAGATAGACGTCATAGGAAG 337
Db 215 GlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLys 234
QY 338 GAGACCGAGCTGAGCTGAAGAGCCATCAGTGTGCACCTCCACCCCTGAGGGCTCTCC 397
Db 235 GluAsnAlaGlyAlaAlaGlyLysAlaIleSerValHisSerThrProGluGlyCysSer 254
QY 398 TCCGCTTGAAGATCATCTTGGAGATTATGCATAAAGAGGCTAAGGACACCAAAACGGCT 457
Db 255 SerAlaCysLysMetIleLeuGluIleMetHisLysGluAlaLysAspThrLysThrAla 274
QY 458 GACGAGTTCCTCCGAGATCTCGCCCAATTAATCTTGTAGGGCGTCTCATTTGGCAAG 517
Db 275 AspGluValProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLys 294
QY 518 GAAGCAGCAACCTGAAGAGGTAGACAGATACCCAGACAAATAATCACATCTCTCG 577
Db 295 GluGlyArgAsnLeuLysLysValGluGlnAspThrGluThrLysIleThrIleSerSer 314
QY 578 TTGAAGACCTTACCTTTACACCTGAGAGGACCATCATCTGTGAAGGGGGCCATCGAG 637
Db 315 LeuGlnAspLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAlaIleGlu 334
QY 638 AATTGTTGCGGCGGACGAGCAATATGAAGAAGTTCGGAGGCGCTATGAGATGAT 697
Db 335 AsnCysCysArgAlaGluGlnGluIleMetLysLysValArgGluAlaTyrGluAsnAsp 354
QY 698 GTGGCTGCCATGAGC-----TCTCACCTGATCCCTGCGCTGAACTGGCTGTAGGT 751
Db 355 ValAlaAlaMetSerLeuGlnSerHisLeuIleProGlyLeuAsnLeuAlaAlaValGly 374
QY 752 CTTTTCAGGCTTATCCAGGCGAGTCCCGCGCTCCAGCAGCGTTCATGGGGTCT 811
Db 375 LeuPheProAlaSerSerAlaValProProProProSerValThrGlyAlaAla 394
QY 812 CCTATAGCTCTTTATGACGGCTCCCGCAGCAGAGATGGTGCAGAGTGTATTATCCCGCC 871
Db 395 ProTyrSerSerPheMetGlnAlaProGluGlnGluMetValGlnValPheIleProAla 414
QY 872 CAGCAGTGGCGGCATCATCGCAAGAGGGGAGCAGCATCAACAGCTCTCCGGTTT 931
Db 415 GlnAlaValGlyAlaIleIleGlyLysGlyGlnHisLysGlnLeuSerArgPhe 434
QY 932 CCCAGCGCTCCATCAAGATTGACCAACCCCGGAAACACCTGACTCCCAAGTTCGTATGGTT 991
Db 435 AlaSerAlaSerIleLysIleAlaProProGluThrProaspSerLysValArgMetVal 454

QY 992 ATCATCACTGGACCGCCAGAGGCCCAATTCAAGGCTCAGGGAAGAATCTATGGCAACTC 1051
Db 455 IleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysLeu 474
QY 1052 AAGGAGGAAACTTCTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACCCACATACGTGTG 1111
Db 475 LysGluGluAsnPheGlyProLysGluGluValLysLeuGluThrHisIleArgVal 494
QY 1112 CCAGCATCAGCAGCTGGCGCGGTCTATTGGCAAGGTGGAAGAACCGTGCAAGCTTCGAG 1171
Db 495 ProAlaSerAlaAlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGln 514
QY 1172 AATTTGACGCGAGCTGAGTGTGTAGTACCAAGACAGACCCCTGTAGTGAACGACCGAG 1231
Db 515 AsnLeuThrAlaAlaGluValValProArgAspGlnThrProAspGluAsnAspGln 534
QY 1232 GTCATCGTGAATAATCATCGGACATTTCTATCCAGTCAAGTGGCTCAACGGAAGATCCGA 1291
Db 535 ValIleValLysIleIleGlyHisPheTyrAlaSerGlnMetAlaGlnArgLysIleArg 554
QY 1292 GACATCCTGGCGCCAGCTTAAGCAGCAGCATCAGAGGACAGAGTAACGAGGCCAGGCA 1351
Db 555 AspIleLeuAlaGlnValLysGlnHisGlnLysGlyGlnSerAsnGlnAlaGlnAla 574
QY 1352 CGGAGGAAG 1360
Db 575 ArgArgLys 577

RESULT 2
US-09-764-864-1114
; Sequence 1114, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1114
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-764-864-1114

Alignment Scores:
Pred. No.: 1.26e-197 Length: 261
Score: 216.00 Matches: 216
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 38.57% Indels: 0
DB: 9 Gaps: 0

US-09-270-437D-5 (1-1708) x US-09-764-864-1114 (1-261)

QY 713 TCTCACTGATCCCTGGCTGNAAGCTGCTGTAGTCTTTCCAGGCTTCATCCAGC 772
Db 46 SerHisLeuIleProGlyLeuAsnLeuAlaValGlyLeuPheProAlaSerSer 65
QY 773 GCAGTCCCGCGCTCCAGCAGCGTGTACTGGGGCTGTCTCCCTATAGTCTCTTATGAG 832
Db 66 AlaValProProProProSerSerValThrGlyAlaAlaProTyrSerSerPheMetGln 85
QY 833 GCTCCCGAGCAGCAGATCGTGCAGGTGTATTATCCCGCCAGCAGTGGGGCCATCATC 892
Db 86 AlaProGluGlnGluMetValGlnValPheIleProAlaGlnAlaValGlyAlaIle 105
QY 893 GGCAGAGAGGGGCGACACATCAACAGCTCTCCCGGTTTGGCAGCGCTCCATCAAGATT 952
Db 106 GlyLysLysGlyGlnHisIleLysGlnLeuSerArgPheAlaSerAlaSerIleLysIle 125
```

QY 953 GCACCACCGGAAACACCTGACTCCAAAGTTCGTATGTTATCATCATCTGACCTGACCCGACAG 1012  
Db 126 AlaProProGluThrProAspSerLysValArgMetValIleIleThrGlyProProGlu 145  
QY 1013 GCCCAATTCAAGGGTCAGGGAGAGATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGT 1072  
Db 146 AlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysLeuLysGluGluAsnPhePheGly 165  
QY 1073 CCCAAGGAGGAAGTGAAGCTGGAGACCCACATAGCTGTGCAGCATAGCAGCTGGCCGG 1132  
Db 166 ProLysGluGluValLysLeuGluThrHisIleArgValProAlaSerAlaAlaGlyArg 185  
QY 1133 GTCATTGGCAAAAGGTGGAAAAACGGTGAACGAGTTCAGATTTTCACGGCAGCTGAGGTG 1192  
Db 186 ValIleGlyLysGlyGlyLysThrValAsnGluLeuGlnAsnLeuThrAlaAlaGluVal 205  
QY 1193 GTAGTACCAAGACGACAGCCCTGATGAGAACGACGACGAGTTCATCGTGAAAAATCATCGGA 1252  
Db 206 ValValProArgAspGlnThrProAspGluAsnAspGlnValIleValLysIleIleGly 225  
QY 1253 CATTCTATGTCAGTCAGATCGCTCAACGGGAAGATCCGAGACATCCTGGCCAGGTTAAG 1312  
Db 226 HisPheTyrAlaSerGlnMetAlaGlnArgLysIleArgAspIleLeuAlaGlnValLys 245  
QY 1313 CAGCAGCATCAGAAGGACAGAGTAACACGAGCCGACGAGCAGGAGGAAG 1360  
Db 246 GlnGlnHisGlnLysGlyGlnSerAsnGlnAlaGlnAlaGArgLys 261

RESULT 3  
US-09-873-637-2  
; Sequence 2, Application US/09873637  
; Patent No. US20020061543A1  
; GENERAL INFORMATION:  
; APPLICANT: Ross, Jeffrey  
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN  
; FILE REFERENCE: 960296.95131  
; CURRENT APPLICATION NUMBER: US/09/873.637  
; CURRENT FILING DATE: 2001-06-04  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 577  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-873-637-2

Alignment Scores:  
Pred. No.: 9,89e-150 Length: 577  
Score: 166.00 Matches: 166  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 29.64% Indels: 0  
DB: 9 Gaps: 0

US-09-270-437D-5 (1-1708) x US-09-873-637-2 (1-577)

QY 215 GTGAGATCCCGCTTCGGCTCCTGGTGCACCCAGATATGGTGCCATTATGGCAAG 274  
Db 194 ValAspIleProLeuArgLeuLeuValProThrGlnTyrValGlyAlaIleIleGlyLys 213  
QY 275 GAGGGGCGCCATCCGCAACATCACAAAACAGACCCAGTCCACAGATAGACGTGCATAGG 334  
Db 214 GluGlyAlaThrIleAArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233  
QY 335 AAGGAGAACCGAGGTGCAGCTGAAAAAGCCATCAGTGTGCACCTCCACCCCTGAGGGCTGC 394  
Db 234 LysGluAsnAlaGlyAlaAlaGluLysAlaIleSerValHisSerThrProGluGlyCys 253  
QY 395 TCCTCCGCTTGAAGATCTTCGAGATTTGAGATTTGATTAAGAGGTGAAGGACACCAAAACG 454  
Db 254 SerSerAlaCysLysMetIleLeuGluMetHisLysGluAlaLysAspThrLysThr 273

QY 455 GCTGACGAGGTTCCCTCCTGAAGATCTCGGCCCATTAATACTTTGTAGGGCGTCTCATTGCG 514  
Db 274 AlaAspGluValProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293  
QY 515 AAGCAAGGACGGAACTCCTGAAGAGGTPAGAGCAAGTACCGAGACAAAAATCACCATCTCC 574  
Db 294 LysGluGlyArgAsnLeuLysValGluGlnAspThrGluThrLysIleThrIleSer 313  
QY 575 TCCTTCCAGACGCTTACCTTTACAACTGAGAGACCATCCTGTGAAGGGGCGCATC 634  
Db 314 SerLeuGlnAspLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAlaIle 333  
QY 635 GAGAATTGTTGACGGCCGAGCAGGAGAAATAATGAAGAAAGTTTCGGGAGGCGCTATGAGAAT 694  
Db 334 GluAsnCysArgAlaGluGlnGluIleMeuLysLysValArgGluAlaTyrGluAsn 353  
QY 695 GATGTGGCTGCCATGAGC 712  
Db 354 AspValAlaAlaMetSer 359

RESULT 4  
US-09-764-864-1532  
; Sequence 1532, Application US/09764864  
; Patent No. US20020132753A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT223  
; CURRENT APPLICATION NUMBER: US/09/764,864  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1792  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1532  
; LENGTH: 250  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (127)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (134)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-764-864-1532

Alignment Scores:  
Pred. No.: 9,44e-102 Length: 250  
Score: 116.00 Matches: 116  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 20.71% Indels: 0  
DB: 9 Gaps: 0

US-09-270-437D-5 (1-1708) x US-09-764-864-1532 (1-250)

QY 1013 GCCCAATTCAAGGCTCAGGGAAGATCTATGCAAACTCAAGAGAGAACTTCTTTGGT 1072  
Db 135 AlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysLeuLysGluGluAsnPheGly 154  
QY 1073 CCCAAGGAGGAAGTGAAGCTGGAGACCCACATACGTGTGCCAGCATCAGCAGCTGGCCGG 1132  
Db 155 ProLysGluGluValLysLeuGluThrHisIleAArgValProAlaSerAlaAlaGlyArg 174  
QY 1133 GTCATTGGCAAAAGGTGGAAAAACGGTGAACGAGTTCGCAAAATTCACGGCAGCTGAGGTG 1192  
Db 175 ValIleGlyLysGlyLysThrValAsnGluLeuGlnAsnLeuThrAlaAlaGluVal 194  
QY 1193 GTAGTACCAAGACGACAGCCCTGATGAGAACGACGAGGTTCATCGTGAANAATCATCGGA 1252  
Db 195 ValValProArgAspGlnThrProAspGluAsnAspGlnValIleValLysIleIleGly 214  
QY 1253 CATTTCATGCCAGTCAGATGCTCAACGGAAGATCCGAGACATCCTGGCCCGGTTAAG 1312

```
Db 215 HisPheTyrAlaSerGlnMetAlaGlnArgLysIleArgPheLeuAlaGlnValLys 234
QY 1313 CAGACGATCAGAGGACAGAGTAAACAGGCCCGCAGCAGGAGGAG 1360
Db 235 GlnGlnHisGlnLysGlyGlnSerAsnGlnAlaGlnAlaArgArgLys 250

RESULT 5
US-09-873-637-18
; Sequence 18, Application US/09873637
; Patent No. US20020061543A1
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/873,637
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-873-637-18

Alignment Scores:
Pred. No.: 1.97e-37 Length: 49
Score: 49.00 Matches: 49
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.75% Indels: 0
DB: 9 Gaps: 0

US-09-270-437D-5 (1-1708) x US-09-873-637-18 (1-49)
QY 476 ATCTGGCCCCATAATACTTTGTAGGCGTCTCATTCGCAAGGAGGACGGAACCTGAAG 535
Db 1 IleuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLys 20
QY 536 AAGGTAGACGACGATACCGACGACAAAAATCACCATCTCCTCGTTGCAAGACCTTACCCCTT 595
Db 21 LysValGluGlnAspThrGluThrLysIleThrIleSerSerLeuGlnAspLeuThrLeu 40
QY 596 TACAACTCTGAGAGGACCATCAGTGTG 622
Db 41 TyrAsnProGluArgThrIleThrVal 49

RESULT 6
US-09-873-637-20
; Sequence 20, Application US/09873637
; Patent No. US20020061543A1
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/873,637
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-873-637-20

Alignment Scores:
Pred. No.: 1.8e-36 Length: 48
Score: 48.00 Matches: 48
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.57% Indels: 0
```

```
DB: 9 Gaps: 0
US-09-270-437D-5 (1-1708) x US-09-873-637-20 (1-48)
QY 1103 ATACGTGTGCCAGCATCAGCAGCTGCGCGGTGCATTGCAAGGTGGAAAAACGGTGAAC 1162
Db 1 IleArgValProAlaSerAlaAlaGlyArgValIleGlyLysGlyGlyThrValAsn 20
QY 1163 GAGTTGCAGAAATTCACGCGCAGCTGAGGTGGTAGTACCAAGAGACACACCCCTGATGAG 1222
Db 21 GluLeuGlnAsnLeuThrAlaAlaGluValValProArgAspGlnThrProAspGlu 40
QY 1223 AACGACCAAGGTGCATCTGAAAATC 1246
Db 41 AsnAspGlnValIleValLysIle 48

RESULT 7
US-09-873-637-17
; Sequence 17, Application US/09873637
; Patent No. US20020061543A1
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/873,637
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-873-637-17

Alignment Scores:
Pred. No.: 1.64e-35 Length: 47
Score: 47.00 Matches: 47
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.39% Indels: 0
DB: 9 Gaps: 0

US-09-270-437D-5 (1-1708) x US-09-873-637-17 (1-47)
QY 233 CTCCTGGTGGCCACCACCATATGTGGTCCCATTTATGCAAGGAGGCGGCCACCATCCGC 292
Db 1 LeuLeuValProThrGlnTyrValGlyAlaIleIleGlyLysGlyGlyAlaThrIleArg 20
QY 293 AACATCACAAAAACAGACCCAGTCCCAAGATAGACGTGCATAGCAAGGAGGAGAACGCAAGTCA 352
Db 21 AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGlyAla 40
QY 353 GCTGAAAAAGCCATCAGTGTG 373
Db 41 AlaGluLysAlaIleSerVal 47

RESULT 8
US-09-873-637-19
; Sequence 19, Application US/09873637
; Patent No. US20020061543A1
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/873,637
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 47
; TYPE: PRT
```



```
; ORGANISM: Mus musculus
US-09-873-637-19

Alignment Scores:
Pred. No.: 1,35e-33 Length: 47
Score: 45.00 Matches: 45
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.04% Indels: 0
DB: 9 Gaps: 0

US-09-270-437D-5 (1-1708) x US-09-873-637-19 (1-47)
QY 857 GTGTTATCCCGCCGCGCAGTGGCGCATTCATCCGACAGAGGGCGACATCAAA 916
Db 1 ValPheileProAlaGlnAlaValAlaIleGlyLysGlyGlnHisIleLys 20
QY 917 CAGCTCTCCCGGTTTGGCAGCGCTCCATCAAGATTGCACCCCGAACAACCTGATCC 976
Db 21 GinLeuSerArgPheAlaSerAlaSerIleLysIleAlaProGluThrProAspSer 40
QY 977 AAAGTTCGTATGGTT 991
Db 41 LysValArgMetVal 45

RESULT 9
US-10-117-982-472
; Sequence 472, Application US/10117982
; Publication No. US20030138438A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Mericle, Barbara
; APPLICANT: Spies, Gregory A.
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C18
; CURRENT APPLICATION NUMBER: US/10/117,982
; CURRENT FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 484
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 472
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-117-982-472

Alignment Scores:
Pred. No.: 5,48e-25 Length: 70
Score: 36.00 Matches: 36
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.43% Indels: 0
DB: 14 Gaps: 0

US-09-270-437D-5 (1-1708) x US-10-117-982-472 (1-70)
QY 254 GTGGTGCCATTATTGGCAAGAGGGGCCACCATCCGCAACATCAACAAACAGACCCAG 313
Db 18 ValGlyAlaIleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGln 37
QY 314 TCCAGATAGACGTGCATAGGAAGAGAGACCGACGTGCAGCTGAAAAA 361
Db 38 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 53

RESULT 10
US-09-270-437D-5 (1-1708) x US-09-873-637-19 (1-47)
QY 857 GTGTTATCCCGCCGCGCAGTGGCGCATTCATCCGACAGAGGGCGACATCAAA 916
Db 1 ValPheileProAlaGlnAlaValAlaIleGlyLysGlyGlnHisIleLys 20
QY 917 CAGCTCTCCCGGTTTGGCAGCGCTCCATCAAGATTGCACCCCGAACAACCTGATCC 976
Db 21 GinLeuSerArgPheAlaSerAlaSerIleLysIleAlaProGluThrProAspSer 40
QY 977 AAAGTTCGTATGGTT 991
Db 41 LysValArgMetVal 45

RESULT 9
US-10-117-982-472
; Sequence 472, Application US/10117982
; Publication No. US20030138438A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Mericle, Barbara
; APPLICANT: Spies, Gregory A.
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C18
; CURRENT APPLICATION NUMBER: US/10/117,982
; CURRENT FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 484
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 472
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-117-982-472

Alignment Scores:
Pred. No.: 5,48e-25 Length: 70
Score: 36.00 Matches: 36
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.43% Indels: 0
DB: 14 Gaps: 0

US-09-270-437D-5 (1-1708) x US-10-117-982-472 (1-70)
QY 254 GTGGTGCCATTATTGGCAAGAGGGGCCACCATCCGCAACATCAACAAACAGACCCAG 313
Db 18 ValGlyAlaIleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGln 37
QY 314 TCCAGATAGACGTGCATAGGAAGAGAGACCGACGTGCAGCTGAAAAA 361
Db 38 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 53

RESULT 10
US-09-270-437D-5 (1-1708) x US-10-313-986-472 (1-70)
QY 254 GTGGTGCCATTATTGGCAAGAGGGGCCACCATCCGCAACATCAACAAACAGACCCAG 313
Db 18 ValGlyAlaIleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGln 37
QY 314 TCCAGATAGACGTGCATAGGAAGAGAGACCGACGTGCAGCTGAAAAA 361
Db 38 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 53

RESULT 11
US-10-408-765A-2088
; Sequence 2088, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2088
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2088

Alignment Scores:
Pred. No.: 4,41e-25 Length: 422
Score: 36.00 Matches: 36
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.43% Indels: 0
DB: 16 Gaps: 0
```

US-09-270-437D-5 (1-1708) x US-10-408-765A-2088 (1-422)

QY 254 GTGGGTGCATTATTGGCAAGAGGGGGCCACCATCCGCAACATCACAACACAGACCCAG 313  
Db 50 ValGlyAlaIleIleGlyLySgluGlyAlaThrIleArgAsnIleThrLysGlnThrGln 69  
QY 314 TCCAAGATAGACGTGCATAGGAAGAGAGACGCGAGTGCAGCTGAAAAA 361  
Db 70 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 85

## RESULT 12

US-09-735-705-176  
; Sequence 176, Application US/09735705  
; Patent No. US20020052329A1

## GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong  
; APPLICANT: Pan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Fanger, Neil

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: 210121.455C14

; CURRENT APPLICATION NUMBER: US/09/735,705

; CURRENT FILING DATE: 2000-12-12

; NUMBER OF SEQ ID NOS: 419

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 176

; LENGTH: 579

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-735-705-176

## Alignment Scores:

Pred. No.: 4,248-25 Length: 579  
Score: 36.00 Matches: 36  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.43% Indels: 0  
DB: 9 Gaps: 0

US-09-270-437D-5 (1-1708) x US-09-735-705-176 (1-579)

QY 254 GTGGGTGCATTATTGGCAAGAGGGGGCCACCATCCGCAACATCACAACACAGACCCAG 313  
Db 207 ValGlyAlaIleIleGlyLySgluGlyAlaThrIleArgAsnIleThrLysGlnThrGln 226

QY 314 TCCAAGATAGACGTGCATAGGAAGAGAGACGCGAGTGCAGCTGAAAAA 361

Db 227 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 242

## RESULT 13

US-09-735-705-348

; Sequence 348, Application US/09735705  
; Patent No. US20020052329A1

## GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Henderson, Robert A.

; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Fanger, Neil  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C14  
; CURRENT APPLICATION NUMBER: US/09/735,705  
; CURRENT FILING DATE: 2000-12-12  
; NUMBER OF SEQ ID NOS: 419  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 348  
; LENGTH: 579

; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-735-705-348

## Alignment Scores:

Pred. No.: 4,248-25 Length: 579  
Score: 36.00 Matches: 36  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.43% Indels: 0  
DB: 9 Gaps: 0

US-09-270-437D-5 (1-1708) x US-09-735-705-348 (1-579)

QY 254 GTGGGTGCATTATTGGCAAGAGGGGGCCACCATCCGCAACATCACAACACAGACCCAG 313

Db 207 ValGlyAlaIleIleGlyLySgluGlyAlaThrIleArgAsnIleThrLysGlnThrGln 226

QY 314 TCCAAGATAGACGTGCATAGGAAGAGAGACGCGAGTGCAGCTGAAAAA 361

Db 227 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 242

## RESULT 14

US-09-850-716A-176  
; Sequence 176, Application US/09850716A  
; Patent No. US20020115139A1

## GENERAL INFORMATION:

; APPLICANT: Kalos, Michael D.  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Retter, Marc W.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C15  
; CURRENT APPLICATION NUMBER: US/09/850,716A  
; CURRENT FILING DATE: 2001-05-07

; NUMBER OF SEQ ID NOS: 440

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 176

; LENGTH: 579

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-850-716A-176

## Alignment Scores:

Pred. No.: 4,248-25 Length: 579  
Score: 36.00 Matches: 36  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.43% Indels: 0  
DB: 9 Gaps: 0

US-09-270-437D-5 (1-1708) x US-09-850-716A-176 (1-579)

QY 254 GTGGGTGCATTATTGGCAAGAGGGGGCCACCATCCGCAACATCACAACACAGACCCAG 313

Db 207 ValGlyAlaIleIleGlyLySgluGlyAlaThrIleArgAsnIleThrLysGlnThrGln 226

QY 314 TCCAAGATAGACGTGCATAGGAAGAGAGACGCGAGTGCAGCTGAAAAA 361

Db 227 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 242

## RESULT 15

US-09-850-716A-348  
; Sequence 348 Application US/09850716A  
; Patent No. US20020115139A1  
; GENERAL INFORMATION:  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Retter, Marc W.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C15  
; CURRENT APPLICATION NUMBER: US/09/850,716A  
; CURRENT FILING DATE: 2001-05-07  
; NUMBER OF SEQ ID NOS: 440  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 348  
; LENGTH: 579  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-850-716A-348

Alignment Scores:  
Pred. No.: 4.24e-25 Length: 579  
Score: 36.00 Matches: 36  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.43% Indels: 0  
DB: 9 Gaps: 0

US-09-270-437D-5 (1-1708) x US-09-850-716A-348 (1-579)

|    |     |  |     |
|----|-----|--|-----|
| Qy | 254 | GTGGTGCCCATTTGGCAGAGGGGGCCACCATCCGCAACATCAAAACAGACCCAG       | 313 |
| Db | 207 | ValGlyAlaIleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGln | 226 |
| Qy | 314 | TCCAAGATAGCGTCATAGGAGAGAACCGCAGGTGCAGCTGAAAAA                | 361 |
| Db | 227 | SerLysIleAspValHisArgLysGluAsnIleGlyAlaAlaGluLys             | 242 |

Search completed: July 13, 2004, 12:49:55  
Job time : 94 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: July 13, 2004, 12:23:53 ; Search time 88 seconds

(without alignments)  
12247.843 Million cell updates/sec

Title: US-09-270-437D-5

Perfect score: 560

Sequence: 1 agggagcgtccgcacgcgc.....attctcttcaggttttaaaa 1708

Scoring table:

OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1017041 seqs, 315518202 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2033934

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp  
-Q=/cgn2 1/USPTO.spool p/US09270437/runat 13072004 121956 9981/app query.fasta\_1.1863  
-DB=SPTREMBL\_25 -QFMT=fastan -SUFFIX=rspt -MINMATCH=0.1 -ILOOPCU=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=quality -THR\_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc  
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09270437 @CGN 1.146 -runat 13072004 121956 9981 -NCFU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPOPEXT=60 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPOPEXT=60 -DELOP=6 -DELEXT=7

Database :

SPTREMBL\_25:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 225   | 40.2        | 577    | 11    | Q8CGX0 rattus norv |

| ID | Q8CGX0  | PRELIMINARY; | PRT; | 577 | AA. |
|----|---|--------------|------|-----|-----|
| AC | Q8CGX0;   |              |      |     |     |
| DT | 01-MAR-2003 (Tremblrel. 23, Created)  |              |      |     |     |
| DT | 01-MAR-2003 (Tremblrel. 23, Last sequence update)   |              |      |     |     |
| DT | 01-OCT-2003 (Tremblrel. 25, Last annotation update)   |              |      |     |     |
| DE | B-actin zipcode binding protein 1.  |              |      |     |     |
| OS | Rattus norvegicus (Rat)   |              |      |     |     |
| OC | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;   |              |      |     |     |
| OX | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.   |              |      |     |     |
| RN | NCBI_TaxID=10116;   |              |      |     |     |
| RP | [1]   |              |      |     |     |
| RC | SEQUENCE FROM N.A.  |              |      |     |     |
| RA | STRAIN=Sprague-Dawley;  |              |      |     |     |
| RT | Tom T., Singer R.H., Bassell G.J.;  |              |      |     |     |
| RL | "Molecular interactions between r2BP1 and b-actin zipcode required for transport of mRNA and stimulation of spine growth."; |              |      |     |     |
| DR | Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.   |              |      |     |     |
| DR | EMBL; AF541940; AAC16210.1;   |              |      |     |     |
| DR | GO: GO:0003676; F:nucleic acid binding; IEA.  |              |      |     |     |
| DR | InterPro; IPR004087; KH dom.  |              |      |     |     |
| DR | InterPro; IPR004088; KH_type_1.   |              |      |     |     |
| DR | InterPro; IPR000504; RNA_rec_mot.   |              |      |     |     |
| DR | Pfam; PF00013; KH; 4.   |              |      |     |     |

#### ALIGNMENTS

RESULT 1

| ID | Q8CGX0  | PRELIMINARY; | PRT; | 577 | AA. |
|----|---|--------------|------|-----|-----|
| AC | Q8CGX0;   |              |      |     |     |
| DT | 01-MAR-2003 (Tremblrel. 23, Created)  |              |      |     |     |
| DT | 01-MAR-2003 (Tremblrel. 23, Last sequence update)   |              |      |     |     |
| DT | 01-OCT-2003 (Tremblrel. 25, Last annotation update)   |              |      |     |     |
| DE | B-actin zipcode binding protein 1.  |              |      |     |     |
| OS | Rattus norvegicus (Rat)   |              |      |     |     |
| OC | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;   |              |      |     |     |
| OX | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.   |              |      |     |     |
| RN | NCBI_TaxID=10116;   |              |      |     |     |
| RP | [1]   |              |      |     |     |
| RC | SEQUENCE FROM N.A.  |              |      |     |     |
| RA | STRAIN=Sprague-Dawley;  |              |      |     |     |
| RT | Tom T., Singer R.H., Bassell G.J.;  |              |      |     |     |
| RL | "Molecular interactions between r2BP1 and b-actin zipcode required for transport of mRNA and stimulation of spine growth."; |              |      |     |     |
| DR | Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.   |              |      |     |     |
| DR | EMBL; AF541940; AAC16210.1;   |              |      |     |     |
| DR | GO: GO:0003676; F:nucleic acid binding; IEA.  |              |      |     |     |
| DR | InterPro; IPR004087; KH dom.  |              |      |     |     |
| DR | InterPro; IPR004088; KH_type_1.   |              |      |     |     |
| DR | InterPro; IPR000504; RNA_rec_mot.   |              |      |     |     |
| DR | Pfam; PF00013; KH; 4.   |              |      |     |     |

DR Pfam: PF00076; rtm; 2.  
 DR SMART: SM00322; KH; 4.  
 DR SMART: SM00360; RRM; 2.  
 DR PROSITE: PS50084; KH TYPE\_1; 4.  
 DR PROSITE: PS50102; RRM; 2.  
 SQ SEQUENCE 577 AA; 63436 MW; 0647676128FBD1EE CRC64;

Alignment Scores:  
 Pred. No.: 255e-225 Length: 577  
 Score: 225.00 Matches: 225  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 40.18% Indels: 0  
 DB: 11 Gaps: 0

US-09-270-437D-5 (1-1708) x Q8CX0 (1-577)

QY 38 GCCATCATGAAGCTGAATGGCCACAGTTGGAGAACCATGCCCTGAAGTCTCTCATATC 97  
 DB 135 AlalleMetLysLeuAsnGlyHisGlnLeuGluAsnHisAlaLeuLysValSerTyrlle 154  
 QY 98 CCCGATGACGATAGACACAGGACCTCAGATCGGCGCGCGAGGGGCTTGGCTCTCG 157  
 DB 155 ProaspGluGlnIleAlaGlnGlyProGluAsnGlyArgGlyGlyPheGlySerArg 174  
 QY 158 GGTGAGCCGCCCGGCTCACCTGTGGCAGCGGGGCGCCAGCCAGACGACGACGAGTG 217  
 DB 175 GlyGlnProArgGlnGlySerProValAlaAlaGlyAlaProAlaLysGlnGlnVal 194  
 QY 218 GACATCCCGCTTGGCTCTGGTGGCCACAGATATGTTGGTGCCTATTGGCAGGAG 277  
 DB 195 AspIleProLeuArgLeuValProThrGlnTrValGlyAlaIleIleGlyLysGlu 214  
 QY 278 GGGGCCACATCCCAACATCACAAAACAGACACAGCCAGTCCCAAGATAGACGTCATAGAG 337  
 DB 215 GlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLys 234  
 QY 338 GAGAACGCGAGTGCAGTGAAGAAAGCCATCAGTGTGCATCCACCCCTGAGGGCTGCTCC 397  
 DB 235 GluAsnAlaGlyAlaAlaGlyLysAlaIleSerValHisSerThrProGluGlyCysSer 254  
 QY 398 TCCGCTGTAGATGATCTTGGAGATTATGCATAAAGAGCTAAGACACCAAAACGGCT 457  
 DB 255 SerAlaCysLysMetIleLeuGluIleMetHisLysGluAlaLysAspThrLysThrAla 274  
 QY 458 GACGAGTTCCTCCGAGATCCTGGCCCATATAAATCTTGTAGGGCTCTCATTTGGCAAG 517  
 DB 275 AspGluValProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLys 294  
 QY 518 GAAGGACGGAACCTGAAGAGGTAGACGAAGATACCGAGACAAATCACCATCTCTCG 577  
 DB 295 GluGlyArgAsnLeuLysLysValGlnAspThrGluThrLysIleThrIleSerSer 314  
 QY 578 TTGCAAGACCTTACCCCTTTACACCTCGAGAGGACCATCACTGTGAAGGGGCGCATCGAG 637  
 DB 315 LeuGlnAspLeuThrLeuTyAsnProGluArgThrIleThrValLysGlyAlaIleGlu 334  
 QY 638 AATTGTTGCGGGCGGAGGAGAAATATGAAGAAAGTTCGGAGGCGCTATGAGATGAT 697  
 DB 335 AsnCysCysArgAlaGluGlnGluIleMetLysLysValArgGluAlaLysGluAsnAsp 354  
 QY 698 GTGGCTGCATGAGC 712  
 DB 355 ValAlaAlaMetSer 359

RESULT 2  
 Q9NZ18  
 ID Q9NZ18 PRELIMINARY; PRT; 577 AA.  
 AC Q9NZ18;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE mRNA-binding protein CRDBP.

OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Ioannidis P., Trangas T., Dimitriadis E., Samiotaki M.,  
 RA Panoutsakopoulos G., Kyriazoglou I., Voutzoulas S., Tsiapalis C.M.,  
 RA Kittas C., Agnantis N., Pandis N.;  
 RT "Ectopic expression of a KH-domain containing protein, highly  
 RT homologous to both human IMP-1 and mouse CRD-BP, in benign and  
 RT malignant mesenchymal tumors.";  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 4 KH DOMAINS.  
 DR EMBL; AF198254; AAF37203.1; -.  
 DR HSSP; F11940; 1CVJ.  
 DR GO; GO:0003676; F:nucleic acid binding; IEA.  
 DR InterPro; IPR004087; KH dom.  
 DR InterPro; IPR004088; KH type\_1.  
 DR InterPro; IPR00504; RNA\_rec\_mot.  
 DR Pfam; PF00013; KH; 4.  
 DR Pfam; PF00076; rtm; 2.  
 DR SMART; SM00322; KH; 4.  
 DR SMART; SM00360; RRM; 2.  
 DR PROSITE; PS50084; KH TYPE\_1; 4.  
 DR PROSITE; PS50102; RRM; 2.  
 SQ SEQUENCE 577 AA; 63456 MW; 0749A060F252D81D CRC64;

Alignment Scores:  
 Pred. No.: 9.69e-212 Length: 577  
 Score: 212.00 Matches: 212  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 37.86% Indels: 0  
 DB: 4 Gaps: 0

US-09-270-437D-5 (1-1708) x Q9NZ18 (1-577)

QY 725 COTGGCCCTGAACCTGGCTGTAGTCTTTTCCAGCTTCATCCAGCGAGTCCCGCG 784  
 DB 366 ProGlyLeuAsnLeuAlaAlaValGlyLeuPheProAlaSerSerAlaValPro 385  
 QY 785 COTCCACAGCGGTACTGGGGCTGCTCCCTATAGCTCTTTATCGAGGCTCCGACGAG 844  
 DB 386 ProProSerSerValThrGlyAlaAlaProTySerSerPheMetGlnAlaProGluGln 405  
 QY 845 GAGATGGTGCAGGTGTTTATCCCGCCAGGCGAGTGGGGCCATCGGCAAGAGGGG 904  
 DB 406 GluMetValGlnValPheIleProAlaGlnAlaValGlyAlaIleIleGlyLysGly 425  
 QY 905 CAGCATCAAAAGCTCTCCGGTTGTCAGCGCTCCATCAAGATTGACACCGCGAA 964  
 DB 426 GlnHisIleLysGlnLeuSerArgPheAlaSerAlaSerIleLysIleAlaProGlu 445  
 QY 965 ACACCTGACTCAAAAGTTCTGTTATCATCTGACCGCCAGAGGCCCAATTCAG 1024  
 DB 446 ThrProAspSerLysValArgMetValIleIleThrGlyProGluAlaGlnPheLys 465  
 QY 1025 GCTCAGGAGAAATCTATGGCAAACTCAAGGAGAGAACTTCTTTGGTCCCAAGAGAA 1084  
 DB 466 AlaGlnGlyArgIleTyTyGlyLysLeuLysGluAsnPhePheGlyProLysGluGlu 485  
 QY 1085 GTGAGCTGGAGACCCACATACGTGTCCAGCATCAGCAGCTGCCGGTCAATGGCAA 1144  
 DB 486 ValLysLeuGluThrHisIleArgValProAlaSerAlaAlaGlyArgValIleGlyLys 505  
 QY 1145 GGTGAGAAACCGTGAACAGTTCAGAAATTTGACCGCAGCTGAGGTGCTAGTACCAAGA 1204  
 DB 506 GlyGlyLysThrValAsnGluLeuGlnAsnLeuThrAlaAlaGluValValProArg 525  
 QY 1205 GACGAGCCCTGATGAGAACGACGAGTCTACCTGAAATCATCGACATTTCTATGCC 1264  
 DB 526 AspGlnThrProAspGluAsnAspGlnValIleValLysIleIleGlyHisPheTyAla 545

|          |  |   |              |
|----------|--|---|--------------|
| Db       | 274  | AlaAspGluValProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly  | 293          |
| Qy       | 515  | AAGSAAGGACCGAACCTGAAGAAGGTAGACAGATACCGAGACAAATAACCATCTCC      | 574          |
| Db       | 294  | LysGluGlyArgAsnLeuLysValGluGlnAspThrGluThrLysIleThrIleSer     | 313          |
| Qy       | 575  | TCGTTGCAAGACCTTACCCCTTACACCCCTGAGAGGACCATCACTGTTGAAGGGGCCATC  | 634          |
| Db       | 314  | SerLeuGlnAspLeuThrLeuTyrrAsnProGluArgThrIleThrValLysGlyAlaIle | 333          |
| Qy       | 635  | GAGAATTTGTCAGGGCCGACGACGAGAAATAATGAAGAAGTTCGGGAGGCCCTATCAGAAT | 694          |
| Db       | 334  | GluAsnCysCysArgAlaGluGlnGluIleWetLysValArgGluAlaTyrrGluAsn    | 353          |
| Qy       | 695  | GATGTGCTGCCATGACG   | 712          |
| Db       | 354  | AspValAlaAlaMetSer  | 359          |
| RESULT 4 |  |   |              |
| ID       | O88477   | PRELIMINARY;  | PRT; 577 AA. |
| AC       | O88477;  |   |              |
| DT       | 01-NOV-1998  | (TrEMBLrel. 08, Created)                                      |              |
| DT       | 01-NOV-1998  | (TrEMBLrel. 08, Last sequence update)                         |              |
| DT       | 01-OCT-2003  | (TrEMBLrel. 25, Last annotation update)                       |              |
| DE       | Coding region determinant binding protein.                             |   |              |
| GN       | IGF2BP1 OR CRDBP.  |   |              |
| OS       | Mus musculus (Mouse).  |   |              |
| OC       | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;      |   |              |
| OC       | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.     |   |              |
| OX       | NCBI_TaxID=10090;  |   |              |
| FN       | [1]  |   |              |
| RP       | SEQUENCE FROM N.A.   |   |              |
| EX       | MEDLINE=92217743; PubMed=1559612;                                      |   |              |
| RA       | Bernstein P.L., Herrick D.J., Prokipcak R.D., Ross J.;                 |   |              |
| RT       | "Control of c-myc mRNA half-life in vitro by a protein capable of      |   |              |
| RT       | binding to a coding region stability determinant.";                    |   |              |
| RL       | Genes Dev. 6:642-654 (1992).   |   |              |
| FN       | [2]  |   |              |
| RP       | SEQUENCE FROM N.A.   |   |              |
| EX       | MEDLINE=9415886; PubMed=6114742;                                       |   |              |
| RA       | Herrick D.J., Ross J.;   |   |              |
| RT       | "The half-life of c-myc mRNA in growing and serum-stimulated cells:    |   |              |
| RT       | influence of the coding and 3' untranslated regions and role of        |   |              |
| RT       | ribosome translocation.";  |   |              |
| RL       | Mol. Cell. Biol. 14:2119-2128 (1994).                                  |   |              |
| FN       | [3]  |   |              |
| RP       | SEQUENCE FROM N.A.   |   |              |
| EX       | MEDLINE=9417348; PubMed=8132663;                                       |   |              |
| RA       | Prokipcak R.D., Herrick D.J., Ross J.;                                 |   |              |
| RT       | "Purification and properties of a protein that binds to the C-terminal |   |              |
| RT       | coding region of human c-myc mRNA.";                                   |   |              |
| RL       | J. Biol. Chem. 269:9261-9269 (1994).                                   |   |              |
| FN       | [4]  |   |              |
| RP       | SEQUENCE FROM N.A.   |   |              |
| EX       | MEDLINE=97322234; PubMed=9178588;                                      |   |              |
| RA       | Leeds P., Kren B.T., Boylan J.M., Betz N.A., Steer C.J.,               |   |              |
| RA       | Gruppuso P.A., Ross J.;  |   |              |
| RT       | "Developmental regulation of CRD-BP, an RNA-binding protein that       |   |              |
| RT       | stabilizes c-myc mRNA in vitro.";                                      |   |              |
| RL       | Oncogene 14:1279-1286 (1997).  |   |              |
| FN       | [5]  |   |              |
| RP       | SEQUENCE FROM N.A.   |   |              |
| RA       | Ross J., Prokipcak R.D., Leeds P., Doyle G.A.R., Betz N.A.,            |   |              |
| RA       | Fleisig A.J.;  |   |              |
| RL       | Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.                |   |              |
| FN       | [6]  |   |              |
| RP       | SEQUENCE FROM N.A.   |   |              |
| RC       | STRAIN=C57BL/6J; TISSUE=Embryo;  |   |              |
| EX       | MEDLINE=21085660; PubMed=11217851;                                     |   |              |
| RA       | Kawai J., Shinagawa A., Shibata K., Yoshino M., Ishii Y.,              |   |              |
| RA       | Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,     |   |              |
| RA       | Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,     |   |              |

RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Glissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustineich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Nombarts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
 RA Hayashizaki Y.,  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 CC -!- SIMILARITY: CONTAINS 4 KH DOMAINS.  
 DR EMBL; AF061569; AAC72743.1; -.  
 DR EMBL; AK013940; BAB29071.1; -.  
 DR HSP; P11940; 1CVJ.  
 DR MGD; MGI:1890357; Igf2bp1.  
 DR GO; GO:0003676; F:nucleic acid binding; IEA.  
 DR InterPro; IPR004087; KH dom.  
 DR InterPro; IPR004088; KH type\_1.  
 DR InterPro; IPR005054; RNA\_rec\_mot.  
 DR Pfam; PF00013; KH; 4.  
 DR Pfam; PF00076; rrm; 2.  
 DR SMART; SM00322; KH; 4.  
 DR SMART; SM00322; KH; 4.  
 DR PROSITE; PS50084; KH\_type\_1; 4.  
 DR PROSITE; PS50102; RRM; 2.  
 SQ SEQUENCE 577 AA; 63450 MW; EPBBIAF2FF9F0344 CRC64;

Alignment Scores:  
 Pred. No.: 1,09e-163 Length: 577  
 Score: 166.00 Matches: 166  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 29.64% Indels: 0  
 DB: 11 Caps: 0

US-09-270-437D-5 (1-1708) x O88477 (1-577)

QY 215 GTGACATCCCTCCCTCGCTCCTCGTCCGCCACCAGCATGTGGTGCATATTGGCAAG 274  
 Db 194 ValAspIleProLeuArgLeuValProThrGlnTyrValGlyAlaIleIleGlyLys 213  
 QY 275 GAGGGGGCCCATCCGCAACATCACAAACAGACCCAGTCCAGATAGACGTGCATAGG 334  
 Db 214 GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233  
 QY 335 AAGGAGAACCGAGTGCAGCTGAAAGAACCCATCAGTGTGCATCCACCCCTGAGGGCTGC 394  
 Db 234 LysGluAsnAlaGlyAlaAlaGluLysAlaIleSerValHisSerThrProGluGlyCys 253  
 QY 395 TCCTCCCTGTAGATGATCTGGAGATTATGCATTAAGAGAGCTAAGGACACAAAGC 454  
 Db 254 SerSerAlaCysLysMetIleLeuGluIleValHisLysGluAlaLysAspThrLysThr 273  
 QY 455 GCTGACGAGGTTCCTCCCTGAAGATCCTGGCCCATTAATACTTGTAGGGCGTCTCATTTGC 514  
 Db 274 AlaAspGluValProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293  
 QY 515 AAGGAGAGCGAAGACCTGAGAGAGGTAGACAGATACCGAGACAAAATCACCATCTCC 574  
 Db 294 LysGluGlyArgAsnLeuLysLysValGluGluAspThrGluThrLysIleThrIleSer 313  
 QY 575 TCGTTGCAAGACCTTACCTTTTACACCTTGAGAGGACCATCACTGTGAAGGGGGCCATC 634  
 Db 314 SerLeuGluAspLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAlaIle 333  
 QY 635 GAGAAATTGTTGCGAGGCGGAGCGAGAAATAATGAAGAAAGTTTCGGGAGCGCTATGAGAAT 694

Db 334 GluAsnCysCysArgAlaGluGlnGluIleMetLysValArgGluAlaTyrGluAsn 353  
 QY 695 GATGTGGTGCATGAGC 712  
 Db 354 AspValAlaAlaMetSer 359

RESULT 5  
 Q80US9 PRELIMINARY; PRT; 577 AA.  
 AC Q80US9;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Igf2bp1 protein.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/Sv x 129/Sv-CP; TISSUE=Embryonic stem cells;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udgin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Wozny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kerteman Y., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smalusz D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/Sv x 129/Sv-CP; TISSUE=Embryonic stem cells;  
 RA Strausberg R.;  
 RL Submitted (Apr-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC051679; AAH51679.1; -.  
 DR GO; GO:0003676; F:nucleic acid binding; IEA.  
 DR InterPro; IPR004087; KH dom.  
 DR InterPro; IPR004088; KH type\_1.  
 DR InterPro; IPR005054; RNA\_rec\_mot.  
 DR Pfam; PF00013; KH; 4.  
 DR Pfam; PF00076; rrm; 2.  
 DR SMART; SM00322; KH; 4.  
 DR SMART; SM00322; KH; 4.  
 DR PROSITE; PS50084; KH\_type\_1; 4.  
 DR PROSITE; PS50102; RRM; 2.  
 SQ SEQUENCE 577 AA; 63378 MW; D439F544257DA3CE CRC64;

Alignment Scores:  
 Pred. No.: 2,03e-110 Length: 577  
 Score: 115.00 Matches: 115  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 20.54% Indels: 0  
 DB: 11 Gaps: 0

US-09-270-437D-5 (1-1708) x Q80US9 (1-577)

QY 995 ATCACTGACCGCAGAGCGCCCAATTCAAGCTCAGGAGAGATCTATGCAAACTCAAG 1054



```

Db 456 IleThrGlyProProGluAlaGlnPheIysAlaGlnGlyArgIleTyrGlyLysLeuLys 475
QY 1055 GAGGAGAACTCTTTGGTCCCAAGGAGAACTGAAGCTGGAGACCCACATACATGTCCTCA 1114
Db 476 GluGluAsnPhePheGlyProLysGluValLysLeuGluThrHisIleArgValPro 495
QY 1115 GCATCAGCAGCTGCGCGGTCAATGGCAAGGTCGAAAGGTCGAAAGGTCGACGTTGCAGAA 1174
Db 496 AlaSerAlaAlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsn 515
QY 1175 TTGACGGCAGCTGAGGTAGTACCAAGAGACACGACCCCTGATGAGAACGACGAGTC 1234
Db 516 LeuThrAlaAlaGluValValProArgAspGlnThrProAspGluAsnAspGlnVal 535
QY 1235 ATCTGAAATCATCGGACATTTCTATGCCAGTCCAGTGGCTCAACCGAAGATCCGAGAC 1294
Db 536 IleValLysIleIleGlyHisPheTyrAlaSerGlnMetAlaGlnArgLysIleArgAsp 555
QY 1295 ATCTGCGCCAGGTAAAGCAGCAGCATCAGAGGACGAGGAGTAAC 1339
Db 556 IleLeuAlaGlnValLysGlnGlnHisGlnLysGlyGlnSerAsn 570
RESULT 6
O42254 PRELIMINARY; PRT; 576 AA.
AC O42254;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Zipcode-binding protein.
GN ZBP1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97220007; PubMed=9121465;
RA Ross A.F., Oleynikov Y.S., Kislauskis E.H., Taneja K.L., Singer R.H.;
RT "Characterization of a beta-actin mRNA zipcode-binding protein.";
RL Mol. Cell. Biol. 17:2158-2165(1997).
CC -!- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR EMBL; AF026527; AAB82295.1; -.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR004087; KH dom.
DR InterPro; IPR004088; KH type 1.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00013; KH; 4.
DR Pfam; PF00076; rrm; 2.
DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS00084; KH TYPE_1; 4.
DR PROSITE; PS50102; RRM; 2.
DR PROSITE; PS00030; RRM_RNP_1; FALSE NEG.
SQ SEQUENCE 576 AA; 63271 MW; 01AAF2D1D81C8811 CRC64;

Alignment Scores:
Pred. No.: 1e-70 Length: 576
Score: 77.00 Matches: 77
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 13.75% Indels: 0
DB: 13 Gaps: 0

US-09-270-437D-5 (1-1708) x O42254 (1-576)

QY 995 ATCACTGACCGCCAGCGGCCCAATTCAGGCTCAGGAGAACTATGCGAACTCAAG 1054
Db 455 IleThrGlyProProGluAlaGlnPheIysAlaGlnGlyArgIleTyrGlyLysLeuLys 474
QY 1055 GAGGAGAACTCTTTGGTCCCAAGGAGAACTGAAGCTGGAGACCCACATACATGTCCTCA 1114

```

```

Db 475 GluGluAsnPhePheGlyProLysGluValLysLeuGluThrHisIleArgValPro 494
QY 1115 GCATCAGCAGCTGCGCGGTCAATGGCAAGGTCGAAAGGTCGAAAGGTCGACGTTGCAGAA 1174
Db 495 AlaSerAlaAlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsn 514
QY 1175 TTGACGGCAGCTGAGGTAGTACCAAGAGACACGACCCCTGATGAGAAC 1225
Db 515 LeuThrAlaAlaGluValValProArgAspGlnThrProAspGluAsn 531
RESULT 7
O73932 PRELIMINARY; PRT; 594 AA.
AC O73932;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE VGI RNA binding protein variant D.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Havin L., Git A., Elisha Z., Oberman F., Yaniv K.,
RA Pressman Schwartz S., Standart N.M., Yisraeli J.K.;
RL Genes Dev. 0:0-0(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98228351; PubMed=9580341;
RA Deshler J.O., Hightett M.I., Abramson T., Schnapp B.J.;
RT "A highly conserved RNA-binding protein for cytoplasmic mRNA
RT localization in vertebrates.";
RL Curr. Biol. 8:489-496(1998).
CC -!- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR EMBL; AF064634; AAC18598.1; -.
DR EMBL; AF055923; AAC41285.1; -.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR004087; KH dom.
DR InterPro; IPR004088; KH type 1.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00013; KH; 4.
DR Pfam; PF00076; rrm; 2.
DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50084; KH TYPE_1; 4.
DR PROSITE; PS50102; RRM; 2.
SQ SEQUENCE 594 AA; 65643 MW; 54CEA7BFF0856DD6 CRC64;

Alignment Scores:
Pred. No.: 4.01e-33 Length: 594
Score: 41.00 Matches: 41
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 7.32% Indels: 0
DB: 13 Gaps: 0

US-09-270-437D-5 (1-1708) x O73932 (1-594)

QY 983 CGTATGTTATCATCATCGACCGCCAGAGCCCAATTCAGGCTCAGGAGAACTCTAT 1042
Db 465 ArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyr 484
QY 1043 GGCAAACTCAAGGAGAGAACTCTTTGGTCCCAAGGAGAACTGAAGCTGGAGACCCAC 1102
Db 485 GlyLysLeuLysGluGluAsnPhePheGlyProLysGluValLysLeuGluThrHis 504
QY 1103 ATA 1105
Db 505 Ile 505
RESULT 8

```



```
DR GO:0003676; F:nucleic acid binding; IEA.
DR InterPro: IPR004087; KH dom.
DR InterPro: IPR004088; KH_type_1.
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF00013; KH; 4.
DR Pfam: PF00076; rrm; 2.
DR SMART: SM00322; KH; 4.
DR SMART: SM00360; RRM; 2.
DR PROSITE: PS50084; KH TYPE_1; 4.
DR PROSITE: PS50102; RRM; 2.
DR PROSITE: PS50102; RRM; 2.
SQ SEQUENCE 579 AA; 63551 MW; 937E601A95D06B77 CRC64;

Alignment Scores:
Pred. No.: 1,24e-21 Length: 579
Score: 30.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.36% Indels: 0
DB: 11 Gaps: 0

US-09-270-437D-5 (1-1708) x Q8C2J9 (1-579)
QY 254 GTGGTGTCATTATTGGCAAGAGGGGCCACCATCGCAACATCAACAAACAGACCAG 313
Db 207 ValGlyAlaIleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGln 226

QY 314 TCCAGATAGACGTGCATAGAGGAGAAC 343
Db 227 SerLysIleAspValHisArgLysGluAsn 236

RESULT 11
Q8CPN8 PRELIMINARY; PRT; 579 AA.
AC Q8CPN8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 10 days embryo cDNA, RKEN full-length enriched library,
DE clone:2610036B18, full insert sequence (igf2 mRNA-binding protein 3)
DE (Insulin-like growth factor 2, binding protein 3).
OS IGFBP3 OR 2610101NRIK OR MIMP3.
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tonita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher M., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kaniya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RN Nature 409:685-690(2001).
RL [2]
RP SEQUENCE FROM N.A.
RA Mori H., Sakakibara S., Imai T., Nakamura Y., Iijima T., Suzuki A.,
RA Yuasa Y., Takeda M., Okano H.;
```

```
RT "Expression of mouse igf2 mRNA-binding protein 3 and its implications
for the developing central nervous system.";
RL J. Neurosci. Res. 0:0-0(2001).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6; TISSUE=Brain, and Olfactory epithelium;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Krzywinski M.I., Skalska U., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6; TISSUE=Brain, and Olfactory epithelium;
RA Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR EMBL; AK011689; BAB27779.1; -.
DR EMBL; AB046173; BAB19755.1; -.
DR EMBL; BC045138; AAB45138.1; -.
DR EMBL; BC049082; AAB49082.1; -.
DR MGD; MGI:1890359; Igfbp3.
DR GO:0003676; F:nucleic acid binding; IEA.
DR InterPro: IPR004087; KH dom.
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF00013; KH; 4.
DR Pfam: PF00076; rrm; 2.
DR SMART: SM00322; KH; 4.
DR SMART: SM00360; RRM; 2.
DR PROSITE: PS50084; KH_TYPE_1; 4.
DR PROSITE: PS50102; RRM; 2.
SQ SEQUENCE 579 AA; 63574 MW; CABD9A4355B392B7 CRC64;

Alignment Scores:
Pred. No.: 1,24e-21 Length: 579
Score: 30.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.36% Indels: 0
DB: 11 Gaps: 0

US-09-270-437D-5 (1-1708) x Q8CPN8 (1-579)
QY 254 GTGGTGTCATTATTGGCAAGAGGGGCCACCATCGCAACATCAACAAACAGACCAG 313
Db 207 ValGlyAlaIleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGln 226

QY 314 TCCAGATAGACGTGCATAGAGGAGAAC 343
Db 227 SerLysIleAspValHisArgLysGluAsn 236

RESULT 12
Q8PW80 PRELIMINARY; PRT; 582 AA.
ID Q8PW80;
AC Q8PW80;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
```

DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
 DE Vg1 RNA binding protein (Decapentaplegic and Vg-related 1, RNA binding  
 DE protein).  
 GN DVIIRBP.  
 OS Brachydanio rerio (zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Zhang Q., Yaniv K., Oberman F., Wolke U., Git A., Fromer M.,  
 RA Taylor W., Meyer D., Standart N., Raz E., Yisraeli J.K.;  
 RA "Vg1 RBP intracellular distribution and evolutionarily conserved  
 RT expression suggest multiple roles during development.";  
 RL Mech. Dev. 0:0-0(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AB; TISSUE=Body;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Schuchman S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Diatchenko L., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Villalón D.K., Muzny K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W.,  
 RA Richardson S., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smallos D.B., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:15899-15903(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AB; TISSUE=Body;  
 RA Strausberg R.;

RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.

CC -!- SIMILARITY: CONTAINS 4 KH DOMAINS.

DR EMBL; AF161270; AAD45610.1; -.

DR EMBL; BC045873; AAH45873.1; -.

DR ZFIN; ZDB-GENE-000308-1; dvirbp.

DR GO; GO:0003676; F:nucleic acid binding; IEA.

DR InterPro; IPR004087; KH dom.

DR InterPro; IPR004088; KH type 1.

DR InterPro; IPR000504; RNA\_rec\_mot.

DR Pfam; PF00013; KH; 4.

DR Pfam; PF00076; rtm; 2.

DR SMART; SM00322; KH; 4.

DR SMART; SM00360; RRM; 2.

DR PROSITE; PS50084; KH\_type\_1; 4.

DR PROSITE; PS50102; RRM; 2.

SQ SEQUENCE 582 AA; 63351 MW; 9DAB63200681B306 CRC64;

Alignment Scores:

|                        |         |               |     |
|------------------------|---------|---------------|-----|
| Pred. No.:             | 2,3e-15 | Length:       | 582 |
| Score:                 | 24.00   | Matches:      | 24  |
| Percent Similarity:    | 100.00% | Conservative: | 0   |
| Best Local Similarity: | 100.00% | Mismatches:   | 0   |
| Query Match:           | 4.29%   | Indels:       | 0   |
| DB:                    | 13      | Gaps:         | 0   |

US-09-270-437D-5 (1-1708) x Q9PW80 (1-582)

QY 467 CCCCTGAAGTCTCTGGCCCAATAACTTTGTAGGGCGTCTACTTGGCAAGGAAGCAGG 526

Db 277 ProLeuLySILeUAlaHisAsnPhEaVGIArGLeuILeGLySGLuGIaYrG 296  
 QY 527 AACCTGAAGAAG 538  
 Db 297 AsnLeuLySlys 300  
 RESULT 13  
 Q9Y6M1 PRELIMINARY; PRT; 556 AA.  
 AC Q9Y6M1;  
 DT 01-NOV-1999 (TRENBLrel. 12, Created)  
 DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)  
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
 DE Hepatocellular carcinoma autoantigen.  
 GN P62.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99207072; PubMed=10190901;  
 RA Zhang J.Y., Chan E.K., Peng X.X., Tan E.M.;  
 RA "A novel cytoplasmic protein with RNA-binding motifs is an autoantigen  
 RT in human hepatocellular carcinoma.";  
 RL J. Exp. Med. 189:1101-1110(1999).  
 CC -!- SIMILARITY: CONTAINS 4 KH DOMAINS.  
 DR EMBL; AF057352; AAD31596.1; -.

DR GO; GO:0005737; Cytoplasm; TAS.

DR GO; GO:0003723; F:RNA binding; TAS.

DR GO; GO:0007345; P:embryogenesis and morphogenesis; TAS.

DR GO; GO:0009386; P:translational attenuation; TAS.

DR InterPro; IPR004087; KH dom.

DR InterPro; IPR004088; KH\_type\_1.

DR InterPro; IPR000504; RNA\_rec\_mot.

DR Pfam; PF00013; KH; 4.

DR Pfam; PF00076; rtm; 2.

DR SMART; SM00322; KH; 4.

DR SMART; SM00360; RRM; 2.

DR PROSITE; PS50084; KH TYPE 1; 4.

DR PROSITE; PS50102; RRM; 2.

SQ SEQUENCE 556 AA; 61842 MW; 1EDEFB100443DDC4 CRC64;

Alignment Scores:

|                        |          |               |     |
|------------------------|----------|---------------|-----|
| Pred. No.:             | 3.87e-10 | Length:       | 556 |
| Score:                 | 19.00    | Matches:      | 19  |
| Percent Similarity:    | 100.00%  | Conservative: | 0   |
| Best Local Similarity: | 100.00%  | Mismatches:   | 0   |
| Query Match:           | 3.39%    | Indels:       | 0   |
| DB:                    | 4        | Gaps:         | 0   |

US-09-270-437D-5 (1-1708) x Q9Y6M1 (1-556)

QY 983 CGTATGGTATCATCACTGGACCCGAGCCCAATTCAGGCTCAGGGAAGATC 1039

Db 431 ArgMetValIleIleThrGlyProGluAlaGlnPheLysAlaGlnGlyArgIle 449

RESULT 14

Q7TP50

ID Q7TP50 PRELIMINARY; PRT; 169 AA.

AC Q7TP50;

DT 01-OCT-2003 (TRENBLrel. 25, Created)

DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)

DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)

DE Ab2-255.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI\_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RA Xu C.S., Li W.Q., Li Y.C., Yang K.J., Yan H.M., Chang C.F., Zhao L.F.,

RA Ma H., Wang L., Wang S.F., Han H.P., Wang G.P., Chai L.Q., Yuan J.Y.,  
RA Shi J.B., Rahman S., Wang Q.N., Zhang J.B.;  
RT "Liver regeneration after PH.";  
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY325199; AAP92600.1; -;  
SQ SEQUENCE 169 AA; 17743 MW; 19BF6295C1000CBA CRC64;

Alignment Scores:  
Pred. No.: 6.17e-07 Length: 169  
Score: 16.00 Matches: 16  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.86% Indels: 0  
DB: 11 Gaps: 0

US-09-270-437D-5 (1-1708) x Q7TF50 (1-169)

Qy 977 AAAGTTGATGTTATCATCATCTGACCCGACGAGCCCAATTCAG 1024  
Db 59 LysValArgMetValIleIleThrGlyProProGluAlaGlnPheLys 74

RESULT 15

Q7TF9 PRELIMINARY; PRT; 545 AA.  
AC Q7TF9;  
DT 01-OCT-2003 (TREMBlrel. 25, Created)  
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)  
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
DE Hypothetical protein.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6; TISSUE=Brain;  
RX MEDLINE=2238257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Wuzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6; TISSUE=Brain;  
RA Strausberg R.;  
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC054552; AAH54552.1; -;  
KW Hypothetical protein.  
SQ SEQUENCE 545 AA; 59679 MW; 88DD35FF30DF3091 CRC64;

Alignment Scores:  
Pred. No.: 5.85e-06 Length: 545  
Score: 15.00 Matches: 15  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.68% Indels: 0  
DB: 11 Gaps: 0

US-09-270-437D-5 (1-1708) x Q7TF9 (1-545)  
Qy 494 TTTGTAGGCGCTCTCATTTGGCAAGCAAGACGACCTGAGAG 538  
Db 277 PheValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLys 291

Search completed: July 13, 2004, 12:34:38  
Job time : 100 secs



| Result No. | Score | Query |        | DB | ID                  | Description        |
|------------|-------|-------|--------|----|---------------------|--------------------|
|            |       | Match | Length |    |                     |                    |
| 1          | 1708  | 100.0 | 1708   | 3  | US-09-061-709-5     | Sequence 5, Appli  |
| 2          | 1708  | 100.0 | 1708   | 4  | US-09-899-651-5     | Sequence 5, Appli  |
| 3          | 1672  | 97.9  | 1946   | 3  | US-09-061-709-7     | Sequence 7, Appli  |
| 4          | 1672  | 97.9  | 1946   | 4  | US-09-899-651-7     | Sequence 7, Appli  |
| 5          | 31    | 1.8   | 2224   | 3  | US-09-261-855-1     | Sequence 1, Appli  |
| 6          | 24    | 1.4   | 24     | 3  | US-09-261-855-38    | Sequence 38, Appli |
| 7          | 23    | 1.3   | 23     | 3  | US-09-261-855-36    | Sequence 36, Appli |
| 8          | 20    | 1.2   | 1740   | 4  | US-09-643-597-347   | Sequence 347, App  |
| 9          | 20    | 1.2   | 1740   | 4  | US-09-542-615A-347  | Sequence 347, App  |
| 10         | 20    | 1.2   | 1740   | 4  | US-09-606-421B-347  | Sequence 347, App  |
| 11         | 20    | 1.2   | 4159   | 3  | US-09-061-709-4     | Sequence 4, Appli  |
| 12         | 20    | 1.2   | 4159   | 3  | US-09-899-651-4     | Sequence 4, Appli  |
| 13         | 20    | 1.2   | 4181   | 4  | US-09-643-597-175   | Sequence 175, App  |
| 14         | 20    | 1.2   | 4181   | 4  | US-09-480-884A-175  | Sequence 175, App  |
| 15         | 20    | 1.2   | 4181   | 4  | US-09-542-615A-175  | Sequence 175, App  |
| 16         | 20    | 1.2   | 4181   | 4  | US-09-606-421B-175  | Sequence 175, App  |
| 17         | 18    | 1.1   | 324    | 4  | US-09-489-039A-4791 | Sequence 4791, Ap  |
| 18         | 18    | 1.1   | 513    | 4  | US-09-252-991A-7364 | Sequence 7364, Ap  |
| 19         | 18    | 1.1   | 1001   | 4  | US-09-571-311-286   | Sequence 286, App  |
| 20         | 18    | 1.1   | 1059   | 4  | US-09-252-991A-7254 | Sequence 7254, Ap  |
| 21         | 18    | 1.1   | 1233   | 4  | US-09-489-039A-4691 | Sequence 4691, Ap  |
| 22         | 18    | 1.1   | 1408   | 1  | US-08-447-554-3     | Sequence 3, Appli  |
| 23         | 18    | 1.1   | 1408   | 1  | US-08-448-160-3     | Sequence 3, Appli  |
| 24         | 18    | 1.1   | 3263   | 3  | US-08-714-918-20    | Sequence 20, Appli |
| 25         | 18    | 1.1   | 3263   | 3  | US-09-265-315-20    | Sequence 20, Appli |
| 26         | 18    | 1.1   | 3263   | 3  | US-09-265-315-20    | Sequence 20, Appli |
| 27         | 18    | 1.1   | 3263   | 3  | US-09-266-417-20    | Sequence 20, Appli |

Db 301 AAAACAGACCAGTCCAGATGAGCTGCTAGAGAGGAGGACGACGAGTGACGCTGAAA 360  
Qy 361 AGCCATCAGTGTGCACTCCACCCCTGAGGGCTGCTCCCTCGCTTGTAAAGATGATCTTGG 420  
Db 361 AGCCATCAGTGTGCACTCCACCCCTGAGGGCTGCTCCCTCGCTTGTAAAGATGATCTTGG 420  
Qy 421 GATTATGATTAAGAGGCTAGGACACCAAAACGGCTGACGAGGTCCCTCGAGATCCT 480  
Db 421 GATTATGATTAAGAGGCTAGGACACCAAAACGGCTGACGAGGTCCCTCGAGATCCT 480  
Qy 481 GGGCCATTAATCTTTGTAGGCGCTCTCATTTGGCAAGAGGACGGAACCTGAAGAAGGT 540  
Db 481 GGGCCATTAATCTTTGTAGGCGCTCTCATTTGGCAAGAGGACGGAACCTGAAGAAGGT 540  
Qy 541 AGAGCAGATACCCAGACAAAATACCATCTCTCGTTGCAAGACCTTACCTTTTACAA 600  
Db 541 AGAGCAGATACCCAGACAAAATACCATCTCTCGTTGCAAGACCTTACCTTTTACAA 600  
Qy 601 CCCTGAGAGGACCATCACTGTGAAGGGGGCCATCGAGAAATTTGTGCAAGGCGGAGCAGGA 660  
Db 601 CCCTGAGAGGACCATCACTGTGAAGGGGGCCATCGAGAAATTTGTGCAAGGCGGAGCAGGA 660  
Qy 661 AATAATGAAGAAAGTTGCGGAGGCTATGAGAAATGATGTGGCTGCCATGAGCTCTACCT 720  
Db 661 AATAATGAAGAAAGTTGCGGAGGCTATGAGAAATGATGTGGCTGCCATGAGCTCTACCT 720  
Qy 721 GATCCCTGGCTGAACCTGGCTGTGTAGGTCTTTTCCAGCTTCATCCAGCGCAGTCCC 780  
Db 721 GATCCCTGGCTGAACCTGGCTGTGTAGGTCTTTTCCAGCTTCATCCAGCGCAGTCCC 780  
Qy 781 GCCGCTCCAGCAGCGTACTGGGGTGTCTCCATAGCTTCTTTATGAGGCTCCCGA 840  
Db 781 GCCGCTCCAGCAGCGTACTGGGGTGTCTCCATAGCTTCTTTATGAGGCTCCCGA 840  
Qy 841 GCAGGATGTGAGGTGTTATCCCGCCAGGAGTGGGGCCATCATCGGCAAGAA 900  
Db 841 GCAGGATGTGAGGTGTTATCCCGCCAGGAGTGGGGCCATCATCGGCAAGAA 900  
Qy 901 GGGGAGCAGATCAAAACAGCTCTCCCGGTTTGGCAGGCGCTCCATCAAGATTGACACCC 960  
Db 901 GGGGAGCAGATCAAAACAGCTCTCCCGGTTTGGCAGGCGCTCCATCAAGATTGACACCC 960  
Qy 961 CGAACACCTGACTCCAAAGTTCGTATGTTATCATCTGAGGAGCGGCGGAGGCGGCAAT 1020  
Db 961 CGAACACCTGACTCCAAAGTTCGTATGTTATCATCTGAGGAGCGGCGGAGGCGGCAAT 1020  
Qy 1021 CAAGGCTCAGGGAAGATCTATGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGGA 1080  
Db 1021 CAAGGCTCAGGGAAGATCTATGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGGA 1080  
Qy 1081 GGAAGTGAAGTGGAGACCCACATACGTGTGCCAGCATACGAGCTGGCGGGTCAATTGG 1140  
Db 1081 GGAAGTGAAGTGGAGACCCACATACGTGTGCCAGCATACGAGCTGGCGGGTCAATTGG 1140  
Qy 1141 CAAAGGTGGAAGAAACGGTGAACAGTTTGCAGAAATTTGACGGCAGCTGAGTGTAGTACC 1200  
Db 1141 CAAAGGTGGAAGAAACGGTGAACAGTTTGCAGAAATTTGACGGCAGCTGAGTGTAGTACC 1200  
Qy 1201 AAGAGACAGACCCCTGATGAGAAACGACACGAGTTCATCGTAAATCATCGGATTTCTA 1260  
Db 1201 AAGAGACAGACCCCTGATGAGAAACGACACGAGTTCATCGTAAATCATCGGATTTCTA 1260  
Qy 1261 TGCCAGTCAAGTGTCTCAACGGAGATCCGAGACATCTCGGCCAGGTTAAGCAGAGCA 1320  
Db 1261 TGCCAGTCAAGTGTCTCAACGGAGATCCGAGACATCTCGGCCAGGTTAAGCAGAGCA 1320  
Qy 1321 TCAGAGGAGCAGAGTAAACAGGCGCCAGGACGAGGAAAGTACCAGGCCCTCCCTGTCC 1380  
Db 1321 TCAGAGGAGCAGAGTAAACAGGCGCCAGGACGAGGAAAGTACCAGGCCCTCCCTGTCC 1380  
Qy 1381 CTTCAGTCCAGGACCAACAGGCGGAGAAATCGAGAGTGTCTCTCCCGGAGGCTGA 1440

Db 1381 CTTCAGTCCAGGACCAACACGGGAGAAATCGAGAGTGTCTCTCCCGGAGGCGCTGA 1440  
Qy 1441 GAATGAGTGGGAATCCGGGACACNTGGGCGGGCTGTAGATCAGGTTTGCCCACTTGATT 1500  
Db 1441 GAATGAGTGGGAATCCGGGACACNTGGGCGGGCTGTAGATCAGGTTTGCCCACTTGATT 1500  
Qy 1501 GAGAAAGATGTTCCAGTGGAGAACCTCATCTNTCAGCCCCCAACACCCCAATTTGGC 1560  
Db 1501 GAGAAAGATGTTCCAGTGGAGAACCTCATCTNTCAGCCCCCAACACCCCAATTTGGC 1560  
Qy 1561 CCAACACTGTTTCCCTCGGGTGTAGAAATTTAGCGCAAGGCACTTTTAAACGTTGG 1620  
Db 1561 CCAACACTGTTTCCCTCGGGTGTAGAAATTTAGCGCAAGGCACTTTTAAACGTTGG 1620  
Qy 1621 ATTGTTTAAAGAGCTCTCCAGGCCCCCAGAGGGTGGATCACACCTCAGTGGGAGA 1680  
Db 1621 ATTGTTTAAAGAGCTCTCCAGGCCCCCAGAGGGTGGATCACACCTCAGTGGGAGA 1680  
Qy 1681 AAAATAAAATTTCTTCAGGTTTAAAA 1708  
Db 1681 AAAATAAAATTTCTTCAGGTTTAAAA 1708

## RESULT 2

US-09-899-651-5  
; Sequence 5, Application US/09899651  
; Patent No. 6576756  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Yao-Tsung  
; APPLICANT: Gure, Ali  
; APPLICANT: Tsong, Solam  
; APPLICANT: Stockert, Elisabeth  
; APPLICANT: Jager, Elke  
; APPLICANT: Knuth, Alexander  
; APPLICANT: Old, Lloyd J.  
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer  
; TITLE OF INVENTION: Associated  
; TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof  
; FILE REFERENCE: LUD 5538  
; CURRENT APPLICATION NUMBER: US/09/899,651  
; PRIOR FILING DATE: 2001-07-06  
; PRIOR APPLICATION NUMBER: US/09/061,709  
; PRIOR FILING DATE: 1998-04-17  
; NUMBER OF SEQ ID NOS: 8  
; SEQ ID NO 5  
; LENGTH: 1708  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
US-09-899-651-5

Query Match 100.0%; Score 1708; DB 4; Length 1708;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGAGCGTGGCGCACCGCCCGAGTTTACCCCGGGAGGCCATCATGAAGCTGAATGGCA 60  
Db 1 AGGAGCGTGGCGCACCGCCCGAGTTTACCCCGGGAGGCCATCATGAAGCTGAATGGCA 60  
Qy 61 CAGTGTGAGAACCATGTCCTTGAAGTCTCTCATCCCCGATGAGGATAGCAGG 120  
Db 61 CAGTGTGAGAACCATGTCCTTGAAGTCTCTCATCCCCGATGAGGATAGCAGG 120  
Qy 121 ACCTGAGATGGGCGCCGAGGGGCTTTGGCTCTCGGGGTGAGCCCGGAGGCTCACC 180  
Db 121 ACCTGAGATGGGCGCCGAGGGGCTTTGGCTCTCGGGGTGAGCCCGGAGGCTCACC 180  
Qy 181 TGTGACGCGGGGGCCCCAGCAAGCAGCAGCAAGTGGACATCCCCCTTCGGCTCTGGT 240  
Db 181 TGTGACGCGGGGGCCCCAGCAAGCAGCAGCAAGTGGACATCCCCCTTCGGCTCTGGT 240  
Qy 241 GCCACCCAGTATGTGGGTGCCATTTTGGCAAGAGGGGGCCACCATCCGACATCAC 300



Db 241 GCCACCCAGTATGTGGGTCACATTTTGGCAAGAGGGGGCCACCATCCGCAACATCAC 300  
QY 301 AAAACAGACCCAGTCCAAAGATAGACGTGCTCATPAGNAGAGAGAACCGAGGTGACGTGAA 360  
Db 301 AAAACAGACCCAGTCCAAAGATAGACGTGCTCATPAGNAGAGAGAACCGAGGTGACGTGAA 360  
QY 361 AGCCATAGTGTGACTCCACCCCTGAGGGGTGCTCTCCGCTTGTAAAGATGATCTTGG 420  
Db 361 AGCCATAGTGTGACTCCACCCCTGAGGGGTGCTCTCCGCTTGTAAAGATGATCTTGG 420  
QY 421 GATTATGATTAAGAGGGCTAAGGACACCAAAACCGCTGACAGATTTCCCTGGAAGATCCT 480  
Db 421 GATTATGATTAAGAGGGCTAAGGACACCAAAACCGCTGACAGATTTCCCTGGAAGATCCT 480  
QY 481 G3CCCATATTAACCTTTGAGGGGCTCTCATTTGGCAAGGAGGACGGAACTGAGAGAGT 540  
Db 481 G3CCCATATTAACCTTTGAGGGGCTCTCATTTGGCAAGGAGGACGGAACTGAGAGAGT 540  
QY 541 AGAGCAAGATACCGAGACAAAAATCACCATCTCTCTGTTGCAAGACCTTACCCCTTACAA 600  
Db 541 AGAGCAAGATACCGAGACAAAAATCACCATCTCTCTGTTGCAAGACCTTACCCCTTACAA 600  
QY 601 CCCTGAGAGGACCATCATCTGTGAGGGGGCCCATCGAGAAATGTTGAGGGGCGGAGCAGGA 660  
Db 601 CCCTGAGAGGACCATCATCTGTGAGGGGGCCCATCGAGAAATGTTGAGGGGCGGAGCAGGA 660  
QY 661 AATAATGAAGAAAGTTCCGGAGGGCTATGAGAAATGATGTGGCTGCCATGAGCTCTCACCT 720  
Db 661 AATAATGAAGAAAGTTCCGGAGGGCTATGAGAAATGATGTGGCTGCCATGAGCTCTCACCT 720  
QY 721 GATCCCTGGCTGAACTGGTGTGTAGAGTCTTTTCCAGCTTCATCCAGGCGAGTCCC 780  
Db 721 GATCCCTGGCTGAACTGGTGTGTAGAGTCTTTTCCAGCTTCATCCAGGCGAGTCCC 780  
QY 781 GCCGCTCCAGCAGCGTTACTGGGGTCTCCCTATAGTCTCTTTATGAGGCTCCCGA 840  
Db 781 GCCGCTCCAGCAGCGTTACTGGGGTCTCCCTATAGTCTCTTTATGAGGCTCCCGA 840  
QY 841 GCAGGAGTGTGAGGTGTTTATCCCGCCAGGCGAGTGGGGCCCATCATCGCAAGAA 900  
Db 841 GCAGGAGTGTGAGGTGTTTATCCCGCCAGGCGAGTGGGGCCCATCATCGCAAGAA 900  
QY 901 GGGGAGCAGACATCAAGAGCTCTCCCGTTGGCAGGCGCTCCATCAAGATTCAGCACC 960  
Db 901 GGGGAGCAGACATCAAGAGCTCTCCCGTTGGCAGGCGCTCCATCAAGATTCAGCACC 960  
QY 961 CGAAACACCTGACTCCAAAGTTCGTATGTTTATCATCTGAGCGCGCAGAGGCCAAT 1020  
Db 961 CGAAACACCTGACTCCAAAGTTCGTATGTTTATCATCTGAGCGCGCAGAGGCCAAT 1020  
QY 1021 CAAGGCTCAGGGAAGATCTATGGAACCTCAAGGAGAGAACTTCTTTGTTCCCAAGGA 1080  
Db 1021 CAAGGCTCAGGGAAGATCTATGGAACCTCAAGGAGAGAACTTCTTTGTTCCCAAGGA 1080  
QY 1081 GGAAGTGAAGCTGGAGACCCACATACGTGTGCCAGCATCAGCAGCTGGCGGGTCATGG 1140  
Db 1081 GGAAGTGAAGCTGGAGACCCACATACGTGTGCCAGCATCAGCAGCTGGCGGGTCATGG 1140  
QY 1141 CAAGGTTGGAAGAAACCGGTGAAACGAGTTGCAGAAATTCAGCGCAGCTGAGGTGATACC 1200  
Db 1141 CAAGGTTGGAAGAAACCGGTGAAACGAGTTGCAGAAATTCAGCGCAGCTGAGGTGATACC 1200  
QY 1201 AAGAGACAGACCCCTGATGAGACGACAGGTGATCGTGAATATCATCGGACATTTCTA 1260  
Db 1201 AAGAGACAGACCCCTGATGAGACGACAGGTGATCGTGAATATCATCGGACATTTCTA 1260  
QY 1261 TGCCAGTCAAGTGGCTCAACGGAAGATCCCGAGACATCTGGCCAGGTTAAGCAGCAGCA 1320  
Db 1261 TGCCAGTCAAGTGGCTCAACGGAAGATCCCGAGACATCTGGCCAGGTTAAGCAGCAGCA 1320  
QY 1321 TCAGAGGACAGAGTAAACGAGGCGCAGGACCGAGGAAGTGAACAGCCCTCTCTGTTC 1380  
Db 1321 TCAGAGGACAGAGTAAACGAGGCGCAGGACCGAGGAAGTGAACAGCCCTCTCTGTTC 1380

QY 1381 CTTNGAGTCCAGGACAAACAAACGGGCGAGAAATCGAGAGTGTGTCTCTCCCGGAGGCGCTGA 1440  
Db 1381 CTTNGAGTCCAGGACAAACAAACGGGCGAGAAATCGAGAGTGTGTCTCTCCCGGAGGCGCTGA 1440  
QY 1441 GAATAGTGGGAATCCGGGACACNTGGGCGGGCTGTAGATCAGGTTTCCGCCACTTGATT 1500  
Db 1441 GAATAGTGGGAATCCGGGACACNTGGGCGGGCTGTAGATCAGGTTTCCGCCACTTGATT 1500  
QY 1501 GAGAAAGATGTTCCAGTGAGGAAACCCCTGATCTINTCAGCCCCCAAAACACCCACCAATTGGC 1560  
Db 1501 GAGAAAGATGTTCCAGTGAGGAAACCCCTGATCTINTCAGCCCCCAAAACACCCACCAATTGGC 1560  
QY 1561 CCAACACTGNTGCCCTCGGGGTGTCAAGAAATTTAGCGAAGGCACTTTTAAAGCTGG 1620  
Db 1561 CCAACACTGNTGCCCTCGGGGTGTCAAGAAATTTAGCGAAGGCACTTTTAAAGCTGG 1620  
QY 1621 ATTGTTTAAAGAGCTCTCCAGGCCCCCACCAGAGGGTGGATCAGCTCAGTGGGAAGA 1680  
Db 1621 ATTGTTTAAAGAGCTCTCCAGGCCCCCACCAGAGGGTGGATCAGCTCAGTGGGAAGA 1680  
QY 1681 AANAATAAATTTCCCTTCAGGTTTTAAAA 1708  
Db 1681 AANAATAAATTTCCCTTCAGGTTTTAAAA 1708

## RESULT 3

US-09-061-709-7  
; Sequence 7, Application US/09061709B  
; Patent No. 6297364  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Yao-Tseung  
; APPLICANT: Gure, Ali  
; APPLICANT: Teang, Solam  
; APPLICANT: Stockert, Elisabeth  
; APPLICANT: Jeger, Elke  
; APPLICANT: Knuth, Alexander  
; APPLICANT: Old, Lloyd J.  
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated  
; FILE REFERENCE: LUD 5538  
; CURRENT APPLICATION NUMBER: US/09/061,709B  
; CURRENT FILING DATE: 1998-04-17  
; NUMBER OF SEQ ID NOS: 8  
; SEQ ID NO 7  
; LENGTH: 1946  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
US-09-061-709-7

Query Match 97.9%; Score 1672; DB 3; Length 1946;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1672; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 AGCCATCATGAAGCTGAATGGCCACAGTTGGAGAACCATGCCCTGAAGGTCTCTACAT 96  
Db 275 AGCCATCATGAAGCTGAATGGCCACAGTTGGAGAACCATGCCCTGAAGGTCTCTACAT 334  
QY 97 CCCCGATGAGCAGATAGACAGGACCTGAGAAATGGGCCCGAGGGGCTTTGGCTCTCG 156  
Db 335 CCCCGATGAGCAGATAGACAGGACCTGAGAAATGGGCCCGAGGGGCTTTGGCTCTCG 394  
QY 157 GGGTCAGCCCCGCCAGGGGTCTACCTGTGGCAGCGGGGGCCCCAGCAAGCAGCAAGT 216  
Db 395 GGGTCAGCCCCGCCAGGGGTCTACCTGTGGCAGCGGGGGCCCCAGCAAGCAGCAAGT 454  
QY 217 GGACATCCCCCTTCGGCTCTCGTCCACCCAGATGTGGGTGCCATTATTTGGCAAGA 276  
Db 455 GGACATCCCCCTTCGGCTCTCGTCCACCCAGATGTGGGTGCCATTATTTGGCAAGA 514  
QY 277 GGGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCCCAAGATAGAGCTGCATAGGAA 336

Db 515 GGGGCCACCATCCGACATCAAAACAGACCCAGTCCAGATAGACGTGATAGGAA 574  
Qy 337 GGAGAACCCAGGTGACGTGTAAGAACCATCAGTGTGACATCCACCCCTGAGGGCTGCTC 396  
Db 575 GGAGAACCCAGGTGACGTGTAAGAACCATCAGTGTGACATCCACCCCTGAGGGCTGCTC 634  
Qy 397 CTCGGCTGTGAAGATGATCTTGGAGATTATGCAATAAGAGGCTAAGGACACCAAAAGGC 456  
Db 635 CTCGGCTGTGAAGATGATCTTGGAGATTATGCAATAAGAGGCTAAGGACACCAAAAGGC 694  
Qy 457 TCAGAGGTTCCCTTGAAGATCTTGGCCCAATAAATCTTGTAGGGCTCTCATTTGGCAA 516  
Db 695 TCAGAGGTTCCCTTGAAGATCTTGGCCCAATAAATCTTGTAGGGCTCTCATTTGGCAA 754  
Qy 517 GGAAGGCGGACCTGAGAGAGGTAGACAGATACCGAGACAAATATCCATCTCTCTC 576  
Db 755 GGAAGGCGGACCTGAGAGAGGTAGACAGATACCGAGACAAATATCCATCTCTCTC 814  
Qy 577 GTTGCAAGACCTTACCTTTTAAACCTTGAGAGGACCATCATCTGTGAAGGGGGCCATCGA 636  
Db 815 GTTGCAAGACCTTACCTTTTAAACCTTGAGAGGACCATCATCTGTGAAGGGGGCCATCGA 874  
Qy 637 GAATTTGTCAGGGCGGAGCGGAAATTAATGAAGAGTTCGGAGGCTTATGAGATGA 696  
Db 875 GAATTTGTCAGGGCGGAGCGGAAATTAATGAAGAGTTCGGAGGCTTATGAGATGA 934  
Qy 697 TGTGGCTGCCATGAGCTCTACCTGATCCCTGGCTGAACCTGGCTGCTGTAGTCTTTT 756  
Db 935 TGTGGCTGCCATGAGCTCTACCTGATCCCTGGCTGNACTGGCTGCTGTAGTCTTTT 994  
Qy 757 CCCAGCTTCATCCAGCGAGTCCCGCCCTCCAGAGCGTTACTGGGGCTGTCCTTA 816  
Db 995 CCCAGCTTCATCCAGCGAGTCCCGCCCTCCAGAGCGTTACTGGGGCTGTCCTTA 1054  
Qy 817 TAGTCTCTTTATGAGGCTCCCGAGCGAGATGTCAGGTGTTATCCCGCCAGGC 876  
Db 1055 TAGTCTCTTTATGAGGCTCCCGAGCGAGATGTCAGGTGTTATCCCGCCAGGC 1114  
Qy 877 AGTGGGCGCCATCATCGGCAAGAGGGGCGAGCATCAAAAGCTCTCCCGTTTGGCAG 936  
Db 1115 AGTGGGCGCCATCATCGGCAAGAGGGGCGAGCATCAAAAGCTCTCCCGTTTGGCAG 1174  
Qy 937 CGCTTCAATCAGATTGACCAACCGGAAACCTGACTCCAAAGTTCGTATGTTATCAT 996  
Db 1175 CGCTTCAATCAGATTGACCAACCGGAAACCTGACTCCAAAGTTCGTATGTTATCAT 1234  
Qy 997 CACTGGACCGGCGAGGCGCCAAATCAAGGCTCAGGGAAGATCTATGGCAACTCAAGGA 1056  
Db 1235 CACTGGACCGGCGAGGCGCCAAATCAAGGCTCAGGGAAGATCTATGGCAACTCAAGGA 1294  
Qy 1057 GGAGAACTTTTGTGTCACAGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1116  
Db 1295 GGAGAACTTTTGTGTCACAGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1354  
Qy 1117 ATCAGCAGCTGGCGGGCTATTGGCAAGAGTGGAAACCGTGAACAGTTCAGAGATT 1176  
Db 1355 ATCAGCAGCTGGCGGGCTATTGGCAAGAGTGGAAACCGTGAACAGTTCAGAGATT 1414  
Qy 1177 GACGCGACCTGAGTGTGTAGTACCAAGAGACAGACCCCTGATGAGAACGACAGTTCAT 1236  
Db 1415 GACGCGACCTGAGTGTGTAGTACCAAGAGACAGACCCCTGATGAGAACGACAGTTCAT 1474  
Qy 1237 CGTGAATATCATCGAGATTTCTATGCGAGTCAAGTGGCTCAAGAGAGTTCGAGACAT 1296  
Db 1475 CGTGAATATCATCGAGATTTCTATGCGAGTCAAGTGGCTCAAGAGAGTTCGAGACAT 1534  
Qy 1297 CTTGCCCCAGGTTAAGCAGCAGCATCAAGAGGACAGAGTAAACAGGCCAGGACCGAG 1356  
Db 1535 CTTGCCCCAGGTTAAGCAGCAGCATCAAGAGGACAGAGTAAACAGGCCAGGACCGAG 1594  
Qy 1357 GAAGTGCACGCCCT 1416  
Db 1595 GAAGTGCACGCCCT 1654

Qy 1417 GTGTGCTCTCCCGGACGGCCTGAGATGAGTGGGAATCCGGGACACNTGGCGGGCTG 1476  
Db 1655 GTGTGCTCTCCCGGACGGCCTGAGATGAGTGGGAATCCGGGACACNTGGCGGGCTG 1714  
Qy 1477 TAGATCAGGTTTGGCCACTTGTATTGAGAAAGATGTTCCAGTGAGGAACCCCTGATCTNCA 1536  
Db 1715 TAGATCAGGTTTGGCCACTTGTATTGAGAAAGATGTTCCAGTGAGGAACCCCTGATCTNCA 1774  
Qy 1537 GCGCCAAACACCCACCCAAATGGCCCAACACTGTTNGCCCTCGGGGTGTGAGAAATNT 1596  
Db 1775 GCGCCAAACACCCACCCAAATGGCCCAACACTGTTNGCCCTCGGGGTGTGAGAAATNT 1834  
Qy 1597 AGCGAAGGCACTTTTAAACGTGGATTGTTTAAAGAGCTCTCCAGGCCCCCAAGAGG 1656  
Db 1835 AGCGAAGGCACTTTTAAACGTGGATTGTTTAAAGAGCTCTCCAGGCCCCCAAGAGG 1894  
Qy 1657 GTGGATCACACCTCAGTGGGAAGAAATAAATTTCTCTCAGGTTTAAAA 1708  
Db 1895 GTGGATCACACCTCAGTGGGAAGAAATAAATTTCTCTCAGGTTTAAAA 1946

## RESULT 4

US-09-899-651-7  
; Sequence 7, Application US/09899651  
; Patent No. 6576756  
; GENERAL INFORMATION:

; APPLICANT: Chen, Yao-Tseng  
; APPLICANT: Gure, Ali  
; APPLICANT: Tsang, Solam  
; APPLICANT: Stockert, Elisabeth  
; APPLICANT: Jager, Elke  
; APPLICANT: Knuth, Alexander  
; APPLICANT: Old, Lloyd J.

; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer

; TITLE OF INVENTION: Associated

; TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof

; FILE REFERENCE: LUD 5538

; CURRENT APPLICATION NUMBER: US/09/899,651

; CURRENT FILING DATE: 2001-07-06

; PRIOR APPLICATION NUMBER: US/09/061,709

; PRIOR FILING DATE: 1998-04-17

; NUMBER OF SEQ ID NOS: 8

; SEQ ID NO 7

; LENGTH: 1946

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

US-09-899-651-7

Query Match 97.9%; Score 1672; DB 4; Length 1946;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1672; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 37 AGCCATCATGAAGCTCAATGGCCACAGTTCGAGAACCATGCTTGAAGGTTCTCTACAT 96

Db 275 AGCCATCATGAAGCTCAATGGCCACAGTTCGAGAACCATGCTTGAAGGTTCTCTACAT 334

Qy 97 CCCCAGATGAGCAGATAGACAGGACCTGAGNATGGGCGCCGAGGGGCTTTGGCTCTCG 156

Db 335 CCCCAGATGAGCAGATAGACAGGACCTGAGNATGGGCGCCGAGGGGCTTTGGCTCTCG 394

Qy 157 GGGTCAGCCCCCGCAGGGCTCACCTGTGGCAGCGGGGCCCCCAGCAACAGCAGCAAGT 216

Db 395 GGGTCAGCCCCCGCAGGGCTCACCTGTGGCAGCGGGGCCCCCAGCAACAGCAGCAAGT 454

Qy 217 GGACATCCCTCTTCCGCTCTCTGGTCCCAACCCAGTATGTGGTGCATTTTGGCAAGA 276

Db 455 GGACATCCCTCTTCCGCTCTCTGGTCCCAACCCAGTATGTGGTGCATTTTGGCAAGA 514

Qy 277 GGGGGSCACATCCGCAACATCACAAAAACAGACCCAGTCCCAAGATAGACGTCATAGAA 336

Db 515 GGGGGSCACATCCGCAACATCACAAAAACAGACCCAGTCCCAAGATAGACGTCATAGAA 574

QY 337 GGAGAACGAGGTGACGCTGAAAGAACCATGATGTGTCACCTCCACCCCTGAGGGCTGCTC 396  
DB 575 GGAGAACGAGGTGACGCTGAAAGAACCATGATGTGTCACCTCCACCCCTGAGGGCTGCTC 634  
QY 397 CTCGCCCTTTGAGATGATCTTGGAGATTATGATGATGAGAGCTAAGGACACCAAAACGGC 456  
DB 635 CTCGCCCTTTGAGATGATCTTGGAGATTATGATGATGAGAGCTAAGGACACCAAAACGGC 694  
QY 457 TGACGAGGTTCCTCGAAGATCTTGGGCCCAATAAATCTTTGTAGGGGTCTCATTTGGCAA 516  
DB 695 TGACGAGGTTCCTCGAAGATCTTGGGCCCAATAAATCTTTGTAGGGGTCTCATTTGGCAA 754  
QY 517 GGAAGGACGAGACCTGGAAGAGGTAGAGCAAGATACCGAGACCAAAATCACCATCTCTC 576  
DB 755 GGAAGGACGAGACCTGGAAGAGGTAGAGCAAGATACCGAGACCAAAATCACCATCTCTC 814  
QY 577 GTTGAAGACCTTACCTTTTAAACCTGTAGAGGACCATCACTGTGAAGGGGGCCATCGA 636  
DB 815 GTTGAAGACCTTACCTTTTAAACCTGTAGAGGACCATCACTGTGAAGGGGGCCATCGA 874  
QY 637 GAATGTGTCAGGGCCGAGCGAGGAATTAATGAAGAAATTCGGGAGGCTATGAGAAATGA 696  
DB 875 GAATGTGTCAGGGCCGAGCGAGGAATTAATGAAGAAATTCGGGAGGCTATGAGAAATGA 934  
QY 697 TGTGGCTGCCATGAGCTCTACCTGATCCCTGGCTGAACCTGGCTGCTGTAGGTCTTTT 756  
DB 935 TGTGGCTGCCATGAGCTCTACCTGATCCCTGGCTGAACCTGGCTGCTGTAGGTCTTTT 994  
QY 757 CCCAGCTTCATCCAGCGAGTCCCGCCCTCCAGCAGCGTTACTGGGGTGTCTCCCTA 816  
DB 995 CCCAGCTTCATCCAGCGAGTCCCGCCCTCCAGCAGCGTTACTGGGGTGTCTCCCTA 1054  
QY 817 TAGCTCTCTTATGACAGGCTCCGAGCAGGAGATGTCAGGTGTTTATCCCGCCCGAGC 876  
DB 1055 TAGCTCTCTTATGACAGGCTCCGAGCAGGAGATGTCAGGTGTTTATCCCGCCCGAGC 1114  
QY 877 AGTGGGCCCATCATTCGCAAGAGGCGGCGAGCATCAACAGCTCTCCCGTTTGGCAG 936  
DB 1115 AGTGGGCCCATCATTCGCAAGAGGCGGCGAGCATCAACAGCTCTCCCGTTTGGCAG 1174  
QY 937 CGCTCTCATCAAGATTGACACCCGAGCAACCTGACTCCAAAGTTCGTATGTTATCAT 996  
DB 1175 CGCTCTCATCAAGATTGACACCCGAGCAACCTGACTCCAAAGTTCGTATGTTATCAT 1234  
QY 997 CACTGGACCGCAGAGGCCCAATTCAGGCTCAGGGAAGATCTATGCAAACTCAAGGA 1056  
DB 1235 CACTGGACCGCAGAGGCCCAATTCAGGCTCAGGGAAGATCTATGCAAACTCAAGGA 1294  
QY 1057 GGAGAACTTCTTGTTCCTCAGAGGAGGAGTGAAGCTGAGACCCACATATCGTGTGCCAGC 1116  
DB 1295 GGAGAACTTCTTGTTCCTCAGAGGAGGAGTGAAGCTGAGACCCACATATCGTGTGCCAGC 1354  
QY 1117 ATCAGCAGCTGGCGGGTCAATGTCAAAGGTGGAAGAACGGTGAACGAGTTGCAGAAATTT 1176  
DB 1355 ATCAGCAGCTGGCGGGTCAATGTCAAAGGTGGAAGAACGGTGAACGAGTTGCAGAAATTT 1414  
QY 1177 GACGCGAGCTGAGGTGTAGTACCAAGAGACAGACCCCTGATGAGAACCGACAGGTCTAT 1236  
DB 1415 GACGCGAGCTGAGGTGTAGTACCAAGAGACAGACCCCTGATGAGAACCGACAGGTCTAT 1474  
QY 1237 CGTGAATATCATCGGACATTTCTATGCCAGTGCAGTGCCTCAACGGAGATCCGAGACAT 1296  
DB 1475 CGTGAATATCATCGGACATTTCTATGCCAGTGCAGTGCCTCAACGGAGATCCGAGACAT 1534  
QY 1297 CTTGGCCCGAGTTAAGCAGCAGCATCAGAAGGGGACAGATTAACAGGCCCGGACCGAG 1356  
DB 1535 CTTGGCCCGAGTTAAGCAGCAGCATCAGAAGGGGACAGATTAACAGGCCCGGACCGAG 1594  
QY 1357 GAAGTGACAGCCCTCCTCTCCCTTNGAGTTCAGGACAAACACGCGGAGAAATCAGA 1416  
DB 1595 GAAGTGACAGCCCTCCTCTCCCTTNGAGTTCAGGACAAACACGCGGAGAAATCAGA 1654

QY 1417 GTGTGCTCTCCCGGAGGCGCTGAGAAATGAGTGGGAATCGGGACACNTGGGCGGGCTG 1476  
DB 1655 GTGTGCTCTCCCGGAGGCGCTGAGAAATGAGTGGGAATCGGGACACNTGGGCGGGCTG 1714  
QY 1477 TAGATCAGGTTTGCCTCAGCTTATTGAGAAAGATGTTCCAGTGTAGGAAACCTGATCTNTCA 1536  
DB 1715 TAGATCAGGTTTGCCTCAGCTTATTGAGAAAGATGTTCCAGTGTAGGAAACCTGATCTNTCA 1774  
QY 1537 GCCCCAAACACCCACCAATTTGGCCCAACACTGNTTGGCCCTGGGGTGTCAAGAAATTT 1596  
DB 1775 GCCCCAAACACCCACCAATTTGGCCCAACACTGNTTGGCCCTGGGGTGTCAAGAAATTT 1834  
QY 1597 AGCGCAAGGCACTTTTAAACGCTGATGTTTAAAGAAAGCTCTCCAGGCCCCCAAGAGG 1656  
DB 1835 AGCGCAAGGCACTTTTAAACGCTGATGTTTAAAGAAAGCTCTCCAGGCCCCCAAGAGG 1894  
QY 1657 GTGGATCACACCTCAGTGGGAGAAATAAATAATTTCTTTCAGGTTTAAAA 1708  
DB 1895 GTGGATCACACCTCAGTGGGAGAAATAAATAATTTCTTTCAGGTTTAAAA 1946

## RESULT 5

US-09-261-855-1  
; Sequence 1, Application US/09261855A  
; Patent No. 6255055  
; GENERAL INFORMATION:  
; APPLICANT: Ross, Jeffrey  
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN  
; FILE REFERENCE: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE  
; CURRENT FILING DATE: 1999-03-03  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 2224  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-261-855-1

Query Match 1.8%; Score 31; DB 3; Length 2224;

Best Local Similarity 100.0%; Pred. No. 1.1e-05;

Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 TGGAGAACCATGCGCTGGAAGGTCTCTCTACAT 96  
DB 561 TGGAGAACCATGCGCTGGAAGGTCTCTCTACAT 591

## RESULT 6

US-09-261-855-38/c  
; Sequence 38, Application US/09261855A  
; Patent No. 6255055  
; GENERAL INFORMATION:  
; APPLICANT: Ross, Jeffrey  
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN  
; FILE REFERENCE: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE  
; CURRENT FILING DATE: 1999-03-03  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 38  
; LENGTH: 24  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Oligonucleotide primer  
US-09-261-855-38

Query Match 1.4%; Score 24; DB 3; Length 24;

Best Local Similarity 100.0%; Pred. No. 0.032;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 508 CATTGGCAAGGAGGACCGAACCT 531  
|||||  
DB 24 CATTGGCAGGAGGACCGAACCT 1

RESULT 7  
US-09-261-855-36  
; Sequence 36, Application US/09261855A  
; Patent No. 6255055  
; GENERAL INFORMATION:  
; APPLICANT: Ross, Jeffrey  
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN  
; FILE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE  
; FILE REFERENCE: 960296.95131  
; CURRENT APPLICATION NUMBER: US/09/261.855A  
; CURRENT FILING DATE: 1999-03-03  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 36  
; LENGTH: 23  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Oligonucleotide primer  
US-09-261-855-36

Query Match 1.3%; Score 23; DB 3; Length 23;  
Best Local Similarity 100.0%; Pred. No. 0.1;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1178 ACGGACGCTGAGGTGGTAGTACC 1200  
|||||  
DB 1 ACGGACGCTGAGGTGGTAGTACC 23

RESULT 8  
US-09-643-597-347  
; Sequence 347, Application US/09643597  
; Patent No. 6426072  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Hendersen, Robert A.  
; APPLICANT: McNeill, Patricia D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C11  
; CURRENT APPLICATION NUMBER: US/09/643.597  
; CURRENT FILING DATE: 2000-08-21  
; NUMBER OF SEQ ID NOS: 369  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 347  
; LENGTH: 1740  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-643-597-347

Query Match 1.2%; Score 20; DB 4; Length 1740;  
Best Local Similarity 100.0%; Pred. No. 4;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1019 TTCAAGGCTCAGGGAAGAAT 1038  
|||||  
DB 1390 TTCAAGGCTCAGGGAAGAAT 1409

RESULT 9  
US-09-542-615A-347  
; Sequence 347, Application US/09542615A  
; Patent No. 6518256  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy A.  
; APPLICANT: Fanger, Gary R.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY  
; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C8  
; CURRENT APPLICATION NUMBER: US/09/542.615A  
; CURRENT FILING DATE: 2000-04-14  
; NUMBER OF SEQ ID NOS: 350  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 347  
; LENGTH: 1740  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-542-615A-347

Query Match 1.2%; Score 20; DB 4; Length 1740;  
Best Local Similarity 100.0%; Pred. No. 4;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1019 TTCAAGGCTCAGGGAAGAAT 1038  
|||||  
DB 1390 TTCAAGGCTCAGGGAAGAAT 1409

RESULT 10  
US-09-606-421B-347  
; Sequence 347, Application US/09606421B  
; Patent No. 6531315  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C9  
; CURRENT APPLICATION NUMBER: US/09/606.421B  
; CURRENT FILING DATE: 2000-06-28  
; NUMBER OF SEQ ID NOS: 358  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 347  
; LENGTH: 1740  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-606-421B-347

Query Match 1.2%; Score 20; DB 4; Length 1740;  
Best Local Similarity 100.0%; Pred. No. 4;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1019 TTCAAGGCTCAGGGAAGAAT 1038  
|||||  
DB 1390 TTCAAGGCTCAGGGAAGAAT 1409

RESULT 11  
US-09-061-709-4  
; Sequence 4, Application US/09061709B

```

; Patent No. 6297364
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tsang
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated
; TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
; FILE REFERENCE: LUD 5538
; CURRENT APPLICATION NUMBER: US/09/061,709B
; CURRENT FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 4
; LENGTH: 4159
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-061-709-4

Query Match      1.2%; Score 20; DB 3; Length 4159;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1019 TTCAAGGCTCAGGGAAGAT 1038
          |||||
Db      1640 TTCAAGGCTCAGGGAAGAT 1659

RESULT 12
US-09-899-651-4
; Sequence 4, Application US/09899651
; Patent No. 6576756
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tsang
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer
; TITLE OF INVENTION: Associated
; TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
; FILE REFERENCE: LUD 5538
; CURRENT APPLICATION NUMBER: US/09/899,651
; CURRENT FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: US/09/061,709
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 4
; LENGTH: 4159
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-899-651-4

Query Match      1.2%; Score 20; DB 4; Length 4159;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1019 TTCAAGGCTCAGGGAAGAT 1038
          |||||
Db      1640 TTCAAGGCTCAGGGAAGAT 1659

RESULT 13
US-09-643-597-175
; Sequence 175, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 175
; LENGTH: 4181
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (3347)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3502)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3506)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3520)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3538)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3549)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3646)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3940)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3968)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3974)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (4036)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (4056)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (4062)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (4080)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (4088)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (4115)
; OTHER INFORMATION: n=A,T,C or G
; OTHER INFORMATION: n=A,T,C or G
US-09-643-597-175

```

```

Query Match          1.2%; Score 20; DB 4; Length 4181;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1019 TTCAAGGCTCAGGGAAGAAT 1038
Db 1640 TTCAAGGCTCAGGGAAGAAT 1659

RESULT 14
US-09-480-884A-175
; Sequence 175, Application US/09480884A
; Patent No. 6482597
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Hosken, Michael D.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C6
; CURRENT APPLICATION NUMBER: US/09/480.884A
; CURRENT FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 330
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 175
; LENGTH: 4181
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(4181)
; OTHER INFORMATION: n=A,T,C or G
US-09-480-884A-175

Query Match          1.2%; Score 20; DB 4; Length 4181;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1019 TTCAAGGCTCAGGGAAGAAT 1038
Db 1640 TTCAAGGCTCAGGGAAGAAT 1659

RESULT 15
US-09-542-615A-175
; Sequence 175, Application US/09542615A
; Patent No. 6518256
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C8
; CURRENT APPLICATION NUMBER: US/09/542.615A
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 175
; LENGTH: 4181
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (3347)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3502)

```

```

; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3506)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3520)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3538)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3549)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3646)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3940)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3968)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3974)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (4036)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (4056)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (4062)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (4080)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (4088)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (4115)
; OTHER INFORMATION: n=A,T,C or G
US-09-542-615A-175

```

```

Query Match          1.2%; Score 20; DB 4; Length 4181;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1019 TTCAAGGCTCAGGGAAGAAT 1038
Db 1640 TTCAAGGCTCAGGGAAGAAT 1659

```

```

Search completed: July 14, 2004, 03:05:05
Job time : 141 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: July 13, 2004, 12:14:48 ; Search time 19.5 Seconds  
(without alignments)  
9121.623 Million cell updates/sec

Title: US-09-270-437d-5  
Perfect score: 560  
Sequence: 1 agggacgtgcgcacgcgc.....atttccttcagggttttaaaa 1708

Scoring table: OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 141681 seqs, 52070155 residues

Word size: 1

Total number of hits satisfying chosen parameters: 283186

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-Q=/cgn2\_1/USPTO\_epool\_p/US09270437/runat\_13072004\_121955\_9970/app\_query.fasta\_1.1863  
-DB=SwissProt\_42 -OPMT=fastan -SUFFIX=rsb -MINMATCH=0.1 -LOOPL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=quality -THR\_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTPMT=pto  
-NCRM=ext -HEA=SIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09270437\_@CGN\_1\_1\_24 @runat\_13072004\_121955\_9970 -NCPU=6 -ICPU=3  
-NO\_WMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| C 1        | 8     | 1.4         | 166    | 1     | ING_CAMBA   |
| C 2        | 8     | 1.4         | 166    | 1     | ING_LAMGL   |
| C 3        | 8     | 1.4         | 224    | 1     | QCRB_BACTC  |
| C 4        | 8     | 1.4         | 227    | 1     | YVEL_BACSU  |
| C 5        | 8     | 1.4         | 259    | 1     | YCBC_ECOLI  |
| C 6        | 8     | 1.4         | 267    | 1     | RS2_PROMO   |
| C 7        | 8     | 1.4         | 371    | 1     | GP85_BRARE  |
| C 8        | 8     | 1.4         | 473    | 1     | PHDK_NOSK   |
| C 9        | 8     | 1.4         | 512    | 1     | DNF5_MOUSE  |
| C 10       | 8     | 1.4         | 600    | 1     | Y018_MYGE   |
| C 11       | 8     | 1.4         | 646    | 1     | SGL_BOVIN   |
| C 12       | 8     | 1.4         | 739    | 1     | BAC1_MOUSE  |
| C 13       | 8     | 1.4         | 815    | 1     | GYRB_MYXXA  |
| C 14       | 8     | 1.4         | 1030   | 1     | Y018_MYCPN  |
| C 15       | 8     | 1.4         | 1473   | 1     | TOP2_ARATH  |
| C 16       | 8     | 1.4         | 1648   | 1     | VJ9H_YEAST  |
| C 17       | 8     | 1.4         | 2567   | 1     | M18B_HUMAN  |
| C 18       | 8     | 1.4         | 3176   | 1     | CA36_HUMAN  |

|      |   |     |      |   |            |        |             |           |
|------|---|-----|------|---|------------|--------|-------------|-----------|
| 19   | 8 | 1.4 | 4451 | 1 | GRSB_BACBR | P14688 | b           | gramicidi |
| C 20 | 8 | 1.4 | 5035 | 1 | RYRI_PIG   | P16960 | sus         | scrofa    |
| C 21 | 8 | 1.4 | 5037 | 1 | RYRI_RABIT | P11716 | oryctolagus |           |
| C 22 | 8 | 1.4 | 5038 | 1 | RYRI_HUMAN | P21817 | homo sapien |           |
| C 23 | 7 | 1.3 | 55   | 1 | RL32_VIBCH | Q9kgh3 | vibrio chol |           |
| C 24 | 7 | 1.2 | 85   | 1 | RM33_YEAST | P20084 | saccharomyc |           |
| C 25 | 7 | 1.3 | 98   | 1 | FIS_ERWCA  | O52540 | erwinia car |           |
| C 26 | 7 | 1.3 | 98   | 1 | FIS_YERPE  | Q8ZAX8 | yersinia pe |           |
| C 27 | 7 | 1.3 | 101  | 1 | CBGR_CLOAB | P24132 | clostridium |           |
| C 28 | 7 | 1.3 | 110  | 1 | LV20_HUMAN | P01713 | homo sapien |           |
| C 29 | 7 | 1.3 | 113  | 1 | GVK1_HALN1 | P24375 | halobacteri |           |
| C 30 | 7 | 1.3 | 114  | 1 | RLA2_BRUMA | P90703 | brugia mala |           |
| C 31 | 7 | 1.3 | 124  | 1 | VA31_VACCC | P21096 | vaccinia vi |           |
| C 32 | 7 | 1.3 | 124  | 1 | VA31_VACCV | P24760 | vaccinia vi |           |
| C 33 | 7 | 1.3 | 136  | 1 | CL17_HUMAN | Q9NXX1 | homo sapien |           |
| C 34 | 7 | 1.2 | 136  | 1 | FL18_VIBCH | Q9K655 | vibrio chol |           |
| C 35 | 7 | 1.3 | 140  | 1 | VA31_VARV  | P33848 | variola vir |           |
| C 36 | 7 | 1.3 | 149  | 1 | GLPA_PANTR | Q28913 | pan troglod |           |
| C 37 | 7 | 1.2 | 150  | 1 | ME15_SCHPO | Q96WS6 | schizosacch |           |
| C 38 | 7 | 1.2 | 153  | 1 | HS11_SOYBN | P02519 | glycine max |           |
| C 39 | 7 | 1.3 | 164  | 1 | ING_CHICK  | P49708 | gallus gall |           |
| C 40 | 7 | 1.3 | 164  | 1 | ING_CORJA  | O57571 | coturnix co |           |
| C 41 | 7 | 1.3 | 164  | 1 | ING_MELGA  | O57603 | meleagris g |           |
| C 42 | 7 | 1.3 | 164  | 1 | ING_NUMME  | O73915 | numida mele |           |
| C 43 | 7 | 1.3 | 164  | 1 | ING_PHACO  | O57608 | phasianus c |           |
| C 44 | 7 | 1.3 | 165  | 1 | GRPI_ORISA | P25074 | oryza sativ |           |
| C 45 | 7 | 1.3 | 173  | 1 | HA34_BRELC | Q99074 | brenia lact |           |

#### ALIGNMENTS

RESULT 1  
ING\_CAMBA  
ID ING\_CAMBA STANDARD; PRT; 166 AA.  
AC Q865W6;  
DT 15-MAR-2004 (Rel. 43, Created)  
DT 15-MAR-2004 (Rel. 43, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Interferon gamma precursor (IFN-gamma).  
GN IFNG.  
OS Camelus bactrianus (Bactrian camel).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.  
OX NCBI\_Taxid=9837;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Raadan O., Lee S., Yoshida R., Chang K., Ohashi K., Sugimoto C.,  
RT Onuma M.;  
RT "Cloning and sequence analysis of cytokine cDNAs of llama and camel."  
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Produced by lymphocytes activated by specific antigens  
CC or mitogens. IFN-gamma, in addition to having antiviral activity,  
CC has important immunoregulatory functions. It is a potent activator  
CC of macrophages, it has antiproliferative effects on transformed  
CC cells and it can potentiate the antiviral and antitumor effects of  
CC the type I interferons (By similarity).  
CC -!- SUBUNIT: Homodimer (By similarity).  
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).  
CC -!- TISSUE SPECIFICITY: Released primarily from activated T  
CC lymphocytes.  
CC -!- SIMILARITY: Belongs to the type II (or gamma) interferon family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AB107657; BAC75394.1; -.  
CC InterPro; IPR002069; IFN-gamma.  
CC Pfam; PF00714; IFN-gamma; 1.  
DR  
DR

DR ProDom; P0002435; IFN-gamma; 1.  
KW Cytokine; Antiviral; Growth regulation; Glycoprotein; Signal.  
FT SIGNAL 1 20 BY SIMILARITY.  
FT CHAIN 21 166 INTERFERON GAMMA.  
FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 106 106 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 166 AA; 19402 MW; 4547EC4PFC693655 CRC64;

## Alignment Scores:

Pred. No.: 36.5 Length: 166  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.44% Indels: 0  
DB: 1 Gaps: 0

US-09-270-437D-5 (1-1708) x ING\_CAMEA (1-166)

QY 298 GATGTTGGGATGGTGCCCTC 275  
|||||  
DB 44 AspValAlaAspGlyGlyProLeu 51

## RESULT 2

ID ING\_LAMGL STANDARD; PRT; 166 AA.  
AC Q865X1.  
DT 15-MAR-2004 (Rel. 43, Created)  
DT 15-MAR-2004 (Rel. 43, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Interferon gamma precursor (IFN-gamma).  
GN IFNG.  
OS Lama glama (Llama).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Lama.  
OX NCBI\_TaxID=9844;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Raadan O., Lee S., Yoshida R., Chang K., Ohashi K., Sugimoto C.,  
RA Onuma M.;  
RL "Cloning and sequence analysis of cytokine cDNAs of llama and camel.";  
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Produced by lymphocytes activated by specific antigens  
CC or mitogens. IFN-gamma, in addition to having antiviral activity,  
CC has important immunoregulatory functions. It is a potent activator  
CC of macrophages, it has antiproliferative effects on transformed  
CC cells and it can potentiate the antiviral and antitumor effects of  
CC the type I interferons (By similarity).  
CC -!- SUBUNIT: Homodimer (By similarity).  
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).  
CC -!- TISSUE SPECIFICITY: Released primarily from activated T  
CC lymphocytes.  
CC -!- SIMILARITY: Belongs to the type II (or gamma) interferon family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AB107652; BAC75389.1; --  
DR InterPro; IPR002069; IFN-gamma.  
DR Pfam; PF00714; IFN-gamma; 1.  
DR ProDom; P0002435; IFN-gamma; 1.  
KW Cytokine; Antiviral; Growth regulation; Glycoprotein; Signal.  
FT SIGNAL 1 20 BY SIMILARITY.  
FT CHAIN 21 166 INTERFERON GAMMA.  
FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 106 106 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 166 AA; 19475 MW; 23937B814759328F CRC64;

## Alignment Scores:

Pred. No.: 36.5 Length: 166  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.44% Indels: 0  
DB: 1 Gaps: 0

US-09-270-437D-5 (1-1708) x ING\_LAMGL (1-166)

QY 298 GATGTTGGGATGGTGCCCTC 275  
|||||  
DB 44 AspValAlaAspGlyGlyProLeu 51

## RESULT 3

ID QCRB\_BACTC STANDARD; PRT; 224 AA.  
AC Q45658;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Menquinol-cytochrome c reductase cytochrome B subunit.  
GN QCRB.  
OS Bacillus thermodenitrificans.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.  
OX NCBI\_TaxID=33940;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K1041.  
RX MEDLINE=96218169; PubMed=8647852;  
RA Sone N., Teuchiya N., Inoue M., Noguchi S.;  
RT "Bacillus stearothermophilus qcr operon encoding Rieske FeS protein, c  
RT cytochrome b6, and a novel-type cytochrome c1 of quinol-cytochrome c  
RT reductase.";  
RL J. Biol. Chem. 271:12457-12462(1996).  
CC -!- FUNCTION: Component of the menaquinol-cytochrome c reductase  
CC complex.  
CC -!- COFACTOR: Two heme groups which are not covalently bound to the  
CC protein (By similarity).  
CC -!- SUBUNIT: THE MAIN SUBUNITS OF THE MENAQUINONE-CYTOCHROME C COMPLEX  
CC ARE: CYTOCHROME B, THE RIESKE PROTEIN AND A 22/29 kDa CYTOCHROME  
CC B/C SUBUNIT.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY. CORRESPONDS TO THE  
CC AMINO END OF MITOCHONDRIAL CYTOCHROME B.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; D83789; BAA12117.1; --  
DR InterPro; IPR005797; Cytb\_b6\_N.  
DR Pfam; PF00033; Cytochrome\_b\_N; 1.  
DR PROSITE; PS00192; CYTOCHROME\_B\_HEME; 1.  
KW Electron transport; Heme; Transmembrane.  
FT TRANSMEM 37 57 POTENTIAL.  
FT TRANSMEM 96 116 POTENTIAL.  
FT TRANSMEM 126 146 POTENTIAL.  
FT TRANSMEM 195 215 POTENTIAL.  
FT METAL 94 94 IRON (HEME 1 AXIAL LIGAND)  
FT METAL 108 108 IRON (HEME 2 AXIAL LIGAND)  
FT METAL 196 196 IRON (HEME 1 AXIAL LIGAND)  
FT METAL 211 211 IRON (HEME 2 AXIAL LIGAND)  
SQ SEQUENCE 224 AA; 25411 MW; AD05B81F4E7F0518 CRC64;

## Alignment Scores:



```

Pred. No.: 35.1 Length: 224
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.43% Indels: 0
DB: 1 Gaps: 0

US-09-270-437D-5 (1-1708) x QCRB_BACTC (1-224)
QY 786 CTCACGAGCGTACTCGGCGTG 809
DB 200 LeuProAlaAlaLeuLeuGlyLeu 207

RESULT 4
YVEL_BACSU
ID YVEL_BACSU STANDARD; PRT; 227 AA.
AC P71051; O08170;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Putative tyrosine-protein kinase yvel (EC 2.7.1.112).
GN YVEL OR BSU34360.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=97124193; PubMed=8969506;
RA Fabret C., Quantin Y., Chapal N., Guiseppe A., Haiech J., Denizot F.;
RT "Integrated mapping and sequencing of a 115 kb DNA fragment from
RT Bacillus subtilis: sequence analysis of a 21 kb segment containing
RT the sigL locus.";
RL Microbiology 142:3089-3096(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Denizot F.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moser I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brocillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.P., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghm S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Guiseppe G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier P., Vassarotti A.,
RA Viaro A., Wambutt R., Wedler E., Wedler H., Weitzneger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumschein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.";
```

```

RL Nature 390:249-256(1997).
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SIMILARITY: Belongs to the cpsd/capB family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z71928; CAA96490.1; -
DR EMBL; Z94043; CAB08024.1; -
DR EMBL; Z99121; CAB15441.1; -
DR PIR; A70036; A70036.
DR Subtilist; BG11860; yvel.
DR InterPro; IPR005702; EPS synthesis.
DR TrRPfams; TR01007; eps_fam; 1.
KW Hypothetical protein; Transferase; Tyrosine-protein kinase;
KW Complete proteome.
SQ SEQUENCE 227 AA; 24674 MW; 81C8B9D75278FFA9 CRC64;

Alignment Scores:
Pred. No.: 35.1 Length: 227
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.44% Indels: 0
DB: 1 Gaps: 0

US-09-270-437D-5 (1-1708) x YVEL_BACSU (1-227)
QY 1367 CTGCTCACTTCCTCGTGGCTGG 1344
DB 48 LeuValThrSerSerValProGly 55

RESULT 5
YCBC_ECOLI
ID YCBC_ECOLI STANDARD; PRT; 259 AA.
AC P36585; P75846;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein ycbc.
GN YCBC OR B0920 OR Z1267 OR ECS1003.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner P.F., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RL "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kaneai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
```

RT corresponding to the 12.7-28.0 min region on the linkage map.";  
 RL DNA Res. 3:137-155(1996).  
 [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;  
 RX MEDLINE=21074935; PubMed=11206551;  
 RA Perna N.T., Plunkett G., III, Burland V., Mau B., Glaesner J.D.,  
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
 RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
 RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouis K.,  
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
 RA Welch R.A., Blattner F.R.;  
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";  
 RL Nature 409:529-533 (2001).  
 [4]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=O157:H7 / RIMD 0509952;  
 RC MEDLINE=21156231; PubMed=11258796;  
 RX Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
 RA Lida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,  
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;  
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli  
 O157:H7 and genomic comparison with a laboratory strain K-12";  
 RL DNA Res. 8:11-22(2001).  
 [5]  
 RN SEQUENCE OF 1-170 FROM N.A.  
 RP STRAIN=K12 / W3110;  
 RC MEDLINE=94232180; PubMed=7513784;  
 RX Feng J., Yamanaka K., Niki H., Ogura T., Hiraga S.;  
 RA "New killing system controlled by two genes located immediately  
 RT upstream of the mukB gene in Escherichia coli";  
 RL Mol. Gen. Genet. 243:136-147(1994).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AE000194; AAC74006.1; -;  
 DR EMBL; D90730; BAA35666.1; -;  
 DR EMBL; AE005281; AAG55405.1; -;  
 DR EMBL; AF002553; BAB34426.1; -;  
 DR EMBL; D26440; -; NOT\_ANNOTATED\_CDS.  
 DR PIR; A85618; A85618.  
 DR PIR; C90754; C90754.  
 DR PIR; G64831; G64831.  
 DR EcolGene; EG12166; YCB.C.  
 DR InterPro; IPR003948; DUF219.  
 DR Pfam; PF02698; DUF218; 1.  
 KW Hypothetical protein; Complete proteome.  
 FT CONFLICT 89 89 Y -> N (IN REF. 5).  
 FT CONFLICT 150 171 GVPEQITLDLPKDTSEAAA -> ACRASKLSPWTCQKI  
 FT PKKLIQ (IN REF. 5).  
 SQ SEQUENCE 259 AA; 28666 MW; 75AE14C696DAA0C9 CRC64;  
 Alignment Scores:  
 Pred. No.: 34.5 Length: 259  
 Score: 8.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.44% Indels: 0  
 DB: 1 Gaps: 0  
 US-09-270-437D-5 (1-1708) x YCB\_CECOLI (1-259)  
 Qy 479 GGATCTTCAGGGAACCTCGTCAG 456  
 Db 251 GlySerSerGlyGluProArgGln 258  
 |||||  
 |||||

RESULT 6  
 RS2\_DROME RS2\_DROME STANDARD; PRT; 267 AA.  
 ID RS2\_DROME RS2\_DROME STANDARD; PRT; 267 AA.  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE 40S ribosomal protein S2 (Strings of pearls protein).  
 GN SOP OR RPS2 OR CS5920.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OC NCBI\_TaxID=7227;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=95073591; PubMed=7982558;  
 RA Cranton S.E., Laeki F.A.;  
 RT "String of pearls encodes Drosophila ribosomal protein S2, has  
 RL Minute-like characteristics, and is required during oogenesis";  
 RN Genetics 137:1039-1048(1994).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=Berkley;  
 RC MEDLINE=20196006; PubMed=10731132;  
 RX Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Adayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brattier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de la Chapelle A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Flosser C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laske P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J., Yao Q.A.,  
 RA Williams S.M., Woodge F., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu S.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster";  
 RL Science 287:2185-2195(2000).  
 [3]  
 RN SEQUENCE OF 1-241 FROM N.A.  
 RP STRAIN=Canton-S;  
 RC MEDLINE=93181212; PubMed=8441641;  
 RX Bartio R., del Arco A., Cabrera H.L., Arribas C.;  
 RA "Cloning and analysis of the S2 ribosomal protein cDNA from  
 RT Drosophila";  
 RL Nucleic Acids Res. 21:351-351(1993).  
 CC -!- SIMILARITY: Belongs to the S2P family of ribosomal proteins.  
 CC -!- SIMILARITY: Contains 1 S5 DREW domain.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----

DR EMBL; X69120; CAA48872.1; --  
 DR EMBL; AE003626; AAF52822.1; --  
 DR EMBL; U01334; AAC34198.1; --  
 DR EMBL; U01335; AAA87053.1; --  
 DR PIR; S30395; S30395.  
 DR HSP; P02357; 1PKP.  
 DR FlyBase; FBgn0004867; sop.  
 DR InterPro; IPR000851; Ribosomal\_S5.  
 DR InterPro; IPR005324; Ribosomal\_S5\_C.  
 DR InterPro; IPR005711; Ribosomal\_S5\_e/a.  
 DR Pfam; PF00333; Ribosomal\_S5; 1.  
 DR Pfam; PF03719; Ribosomal\_S5\_C; 1.  
 DR TIGRFAMs; TIGR01020; rpsE\_arch; 1.  
 DR PROSITE; PS00585; RIBOSOMAL\_S5; 1.  
 DR PROSITE; PS00881; S5\_DSRBD; 1.  
 KW Ribosomal protein.  
 FT DOMAIN 85 148 S5 DBM.  
 FT CONFLICT 19 20 GG -> PP (IN REF. 3).  
 FT CONFLICT 194 194 K -> R (IN REF. 3).  
 SQ SEQUENCE 267 AA; 28899 MW; ADA22CD28F100743 CRC64;

Alignment Scores:  
 Pred. No.: 34.4 Length: 267  
 Score: 8.00 Matches: 8  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.43% Indels: 0  
 DB: 1 Gaps: 0

US-09-270-437D-5 (1-1708) x RS2\_DROME (1-267)

QY 137 CGAGGGGCTTGGCTCTCGGGGT 160

Db 12 ArgGlyGlyPheGlySerArgGly 19

RESULT 7

ID GP85\_BRARE STANDARD; PRT; 371 AA.

AC Q91919;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Super conserved receptor expressed in brain 2.  
 GN SREB2.  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RX MEDLINE=20294882; PubMed=10833454;  
 RA Katsumoto M., Saito T., Takasaki J., Kamohara M., Sugimoto T.,  
 RA Kobayashi M., Tadokoro M., Matsumoto S., Ohishi T., Furuichi K.;  
 RT "An evolutionarily conserved G-protein coupled receptor family, SREB,  
 RT expressed in the central nervous system."  
 RL Biochem. Biophys. Res. Commun. 272:576-582(2000).  
 CC -!- FUNCTION: Orphan receptor.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).  
 CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----

DR EMBL; AB040805; BAA96651.1; --  
 DR ZFIN; ZDB-GENE-000710-2; sreb2.  
 DR InterPro; IPR000276; GPCR\_Rhodpsn.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00237; GPCRRHODPSN.  
 DR PROSITE; PS00237; G PROTEIN RECEPTOR FL1; FALSE\_NEG.  
 DR PROSITE; PS00262; G\_PROTEIN\_RECEP\_FL2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;  
 KW Multigene family.

FT DOMAIN 1 26 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 27 47 1 (POTENTIAL).  
 FT DOMAIN 48 58 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 59 79 2 (POTENTIAL).  
 FT DOMAIN 80 96 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 97 117 3 (POTENTIAL).  
 FT DOMAIN 118 138 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 139 159 4 (POTENTIAL).  
 FT DOMAIN 160 189 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 190 210 5 (POTENTIAL).  
 FT DOMAIN 211 287 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 288 308 6 (POTENTIAL).  
 FT DOMAIN 309 321 7 (POTENTIAL).  
 FT TRANSMEM 322 342 7 (POTENTIAL).  
 FT DOMAIN 343 371 CYTOPLASMIC (POTENTIAL).  
 FT DISULFID 95 173 BY SIMILARITY.  
 FT CARBOHYD 3 3 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 84 84 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 183 183 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 371 AA; 41954 MW; F6F6175ED3A348C2 CRC64;

Alignment Scores:

Pred. No.: 33 Length: 371  
 Score: 8.00 Matches: 8  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.44% Indels: 0  
 DB: 1 Gaps: 0

US-09-270-437D-5 (1-1708) x GP85\_BRARE (1-371)

QY 802 ACTAACGCTGCGGAGCGCGG 779

Db 264 SerAsnAlaGlyArgArg 271

RESULT 8

PHDK NOCSK

ID PHDK NOCSK STANDARD; PRT; 473 AA.

AC 024723;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Probable 1-hydroxy-2-naphthoate transporter.  
 OS Nocardioides sp. (strain KP7).  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Propionibacterineae; Nocardioidaceae; Nocardioides.  
 OX NCBI\_TaxID=35761;  
 RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=97474276; PubMed=9335300;

RA Iwabuchi T., Harayama S.;  
 RT "Biochemical and genetic characterization of 2-carboxybenzaldehyde  
 RT dehydrogenase, an enzyme involved in phenanthrene degradation by  
 RT Nocardioides sp. strain KP7.";  
 RL J. Bacteriol. 179:6488-6494(1997).  
 CC -!- FUNCTION: PROBABLY INVOLVED IN THE TRANSPORT OF 1-HYDROXY-2-  
 CC NAPHTHOATE.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
 CC (Potential).

```

CC CC -!- SIMILARITY: Belongs to the sugar transporter family.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; AB000735; BAA23264.1; -.
CC CC InterPro; IPR007114; MFS.
CC CC InterPro; IPR005828; Sub-transporter.
CC CC InterPro; IPR005829; Sug-transporter.
CC CC Pfam; PF00083; sugar_tr; 1.
CC CC PROSITE; PS00850; MFS; 1.
CC CC PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
CC CC PROSITE; PS00217; SUGAR_TRANSPORT_2; FALSE_NEG.
CC CC Transprot; Transmembrane; Inner membrane.
CC CC DOMAIN 1 25 CYTOPLASMIC (POTENTIAL).
CC CC TRANSMEM 26 46 1 (POTENTIAL).
CC CC DOMAIN 27 59 EXTRACELLULAR (POTENTIAL).
CC CC TRANSMEM 60 80 2 (POTENTIAL).
CC CC DOMAIN 81 91 CYTOPLASMIC (POTENTIAL).
CC CC TRANSMEM 92 112 3 (POTENTIAL).
CC CC DOMAIN 113 121 EXTRACELLULAR (POTENTIAL).
CC CC TRANSMEM 122 142 4 (POTENTIAL).
CC CC DOMAIN 143 153 CYTOPLASMIC (POTENTIAL).
CC CC TRANSMEM 154 174 5 (POTENTIAL).
CC CC DOMAIN 175 180 EXTRACELLULAR (POTENTIAL).
CC CC TRANSMEM 181 201 6 (POTENTIAL).
CC CC DOMAIN 202 263 CYTOPLASMIC (POTENTIAL).
CC CC TRANSMEM 264 284 7 (POTENTIAL).
CC CC DOMAIN 285 304 EXTRACELLULAR (POTENTIAL).
CC CC TRANSMEM 305 325 8 (POTENTIAL).
CC CC DOMAIN 326 330 CYTOPLASMIC (POTENTIAL).
CC CC TRANSMEM 331 351 9 (POTENTIAL).
CC CC DOMAIN 352 354 EXTRACELLULAR (POTENTIAL).
CC CC TRANSMEM 355 375 10 (POTENTIAL).
CC CC DOMAIN 376 398 CYTOPLASMIC (POTENTIAL).
CC CC TRANSMEM 399 418 11 (POTENTIAL).
CC CC DOMAIN 419 421 EXTRACELLULAR (POTENTIAL).
CC CC TRANSMEM 422 444 12 (POTENTIAL).
CC CC DOMAIN 445 473 CYTOPLASMIC (POTENTIAL).
CC CC SEQUENCE 473 AA; 49109 MW; D6D765D376260D8A CRC64;

Alignment Scores:
Pred. No.: 32 Length: 473
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.44% Indels: 0
DB: 1 Gaps: 0

US-09-270-437D-5 (1-1708) x PHDK_NOSCK (1-473)
QY 1217 CAGGGCTGCTGCTCTGTTGACTA 1194
Db 277 GlnGlySerGlyLeuLeuValLeu 284
RESULT 9
DPN5_MOUSE
ID DFN5_MOUSE STANDARD; PRT; 512 AA.
AC Q922D3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Nonsyndromic hearing impairment protein 5 homolog.
GN DFN55 OR DFN55H.
OS Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC NCBI_TaxID=10090;

```

```

RN RP SEQUENCE FROM N.A.
RC TISSUE=Cochlea;
RX MEDLINE=98442658; PubMed=9771715;
RA Van Laer L., Huizing E.H., Verstreken M., van Zuijlen D.,
RA Wauters J.G., Bosesuyt P.J., Van de Heyning P., McGuirt W.T.,
RA Smith R.J.H., Willems P.J., Legan P.K., Richardson G.P., Van Camp G.;
RT "Nonsyndromic hearing impairment is associated with a mutation in
RT DFN55.";
RL Nat. Genet. 20:194-197(1998).
CC -!- SIMILARITY: BELONGS TO THE DFN5 FAMILY.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; AF073309; AAC69325.1; -.
CC CC MGD; MGI:1889850; Dfn5a.
CC CC InterPro; IPR007677; DFN5A.
CC CC InterPro; IPR007681; Mob1.
CC CC Pfam; PF04598; DFN5A; 1.
CC CC SEQUENCE 512 AA; 56630 MW; 13AFB8627773C4A5 CRC64;

Alignment Scores:
Pred. No.: 31.7 Length: 512
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.44% Indels: 0
DB: 1 Gaps: 0

US-09-270-437D-5 (1-1708) x DFN5_MOUSE (1-512)
QY 616 GATGCTCTCTCAGGGTGTGTAAG 593
Db 279 AspGlyProLeuArgValVallys 286
RESULT 10
Y018_MYCGE
ID Y018_MYCGE STANDARD; PRT; 600 AA.
AC P47264; Q49209; Q49302;
DT 01-FEB-1996 (Rel. 33, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical helicase MG018.
GN MG018.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2097;
RN [1]_TaxID=2097;
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=96026346; PubMed=7569993;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Bitt K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT "The minimal gene complement of Mycoplasma genitalium.";
RL Science 270:397-403(1995).
RN [2]
RP REVISIONS.
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Bitt K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;

```

```

RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 209-309 AND 371-471 FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=94075230; PubMed=8253680;
RA Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III;
RT "A survey of the Mycoplasma genitalium genome by using random
sequencing.";
RL J. Bacteriol. 175:7918-7930 (1993).
CC -!- MISCELLANEOUS; IN M.PNEUMONIAE, A SINGLE ORF SPANS M.GENITALIUM
MG016, MG017 AND MG018.
CC -!- SIMILARITY: Belongs to the SNF2/RAD54 helicase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U39681; AAC71234.1; -.
DR EMBL; U02179; AAD12465.1; -.
DR EMBL; U01757; AAD10571.1; -.
DR PIR; T09676; T09676.
DR TIGR; MG018; -.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR000330; SNF2_N.
DR Pfam; PF00271; helicase_C; 1.
DR Pfam; PF00176; SNF2_N; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICC; 1.
KW Hypothetical protein; Helicase; ATP-binding; Complete proteome.
FT NP_BIND 172 179 ATP (POTENTIAL).
FT SITE 271 274 DEAD BOX.
FT CONFLICT 452 462 D -> S (IN REF. 3).
SQ SEQUENCE 600 AA; 68873 MW; 75A82D909F4A877A CRC64;

Alignment Scores:
Pred. No.: 31 Length: 600
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.43% Indels: 0
DB: 1 Gaps: 0

US-09-270-437D-5 (1-1708) x Y019_MTCGE (1-600)
QY 1172 AATTGACGGCAGCTGAGTGCTA 1195
D 515 AsnLeuThrAlaAlaGluValVal 522

RESULT 11
SGI_BOVIN
ID SGI_BOVIN STANDARD; PRT; 646 AA.
AC P23389; O02707;
DT 01-NOV-1991 (Rel. 20, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Secretogranin I precursor (Sgi) (Chromogranin B) (CgB) [Contains: GAWK
DE peptide; Secretolytin].
GN CHGE.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Adrenal chromaffin;
RX MEDLINE=91223091; PubMed=2025642;
RA Bauer J.W., Fischer-Colbrie R.;

```

```

RT "Primary structure of bovine chromogranin B deduced from cDNA
sequence.";
RL Biochim. Biophys. Acta 1089:124-126 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Adrenal medulla;
RX MEDLINE=97282588; PubMed=9136897;
RA Yoo S.H., Kang Y.K.;
RT "Identification of the secretory vesicle membrane binding region of
chromogranin B.";
RL FEBS Lett. 406:259-262 (1997).
RN [3]
RP SEQUENCE OF 21-646 FROM N.A.
RC TISSUE=Adrenal medulla;
RA Grandy D.K., Jeduc R., Makam H., Flanagan T., Diliberto E.J.,
RA Thomas G., Civelli O., Viveros O.H.;
RL Submitted (OCT-1990) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 634-646.
RC TISSUE=Adrenal chromaffin;
RX MEDLINE=95262699; PubMed=7744058;
RA Strub J.-M., Garcia-Sabione P., Lonnig K., Taupenot L., Hubert P.,
RA van Dorsselaer A., Aunis D., Metz-Boutigue M.-H.;
RT "Processing of chromogranin B in bovine adrenal medulla.
RT Identification of secretolytin, the endogenous C-terminal fragment of
RT residues 614-626 with antibacterial activity.";
RL Eur. J. Biochem. 229:356-368 (1995).
RN [5]
RP CHARACTERIZATION OF SECRETOLYTIN.
RX MEDLINE=96184581; PubMed=8603705;
RA Strub J.M., Hubert P., Nullans G., Aunis D., Metz-Boutigue M.-H.;
RT "Antibacterial activity of secretolytin, a chromogranin B-derived
RT peptide (614-626) is correlated with peptide structure.";
RL FEBS Lett. 379:273-278 (1996).
CC -!- FUNCTION: Secretogranin I is a neuroendocrine secretory granule
CC protein, which may be the precursor for other biologically active
CC peptides. The 16 pairs of basic AA distributed throughout its
CC sequence may be used as proteolytic cleavage sites.
CC -!- FUNCTION: Secretolytin has antibacterial activity.
CC -!- SUBCELLULAR LOCATION: Neuroendocrine and endocrine secretory
CC granules.
CC -!- PTM: O-glycosylated (Probable).
CC -!- SIMILARITY: Belongs to the chromogranin / secretogranin protein
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration -
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X55027; CAA38846.1; -.
DR EMBL; U88551; AAC48720.1; -.
DR EMBL; X55489; CAA39109.1; -.
DR PIR; S15901; S15901.
DR InterPro; IPR001819; Chromogranin_AB.
DR InterPro; IPR001930; Granin.
DR Pfam; PF01271; Granin; 1.
DR PRINTS; PR00659; CHROMOGRANIN.
DR PROSITE; PS00422; GRANINS_1; 1.
DR PROSITE; PS00423; GRANINS_2; 1.
DR Sulfation; Cleavage on pair of basic residues; Signal.
FT SIGNAL 1 20
FT CHAIN 21 646 SECRETOGRANIN I.
FT PEPTIDE 418 484 GAWK PEPTIDE.
FT PEPTIDE 634 646 SECRETOLYTIN.
FT DISULFID 36 57 BY SIMILARITY.
FT MOD_RES 158 158 SULFATION (POTENTIAL).
FT MOD_RES 315 315 SULFATION (BY SIMILARITY).
FT CONFLICT 64 70 N -> S (IN REF. 1).
FT CONFLICT 70 70 N -> D (IN REF. 2).

```

```

FT CONFLICT 93 98 SEAPGL -> FRSPRAS (IN REF. 3).
FT CONFLICT 181 181 T -> M (IN REF. 2).
FT CONFLICT 261 261 H -> R (IN REF. 2).
FT CONFLICT 386 386 P -> R (IN REF. 2).
FT CONFLICT 481 481 H -> L (IN REF. 3).
FT CONFLICT 597 597 M -> V (IN REF. 2).
SQ SEQUENCE 646 AA; 73339 MW; 420DB1178D9E415 CRC64;

Alignment Scores:
Pred. No.: 30.8 Length: 646
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.43% Indels: 0
DB: 1 Gaps: 0

US-09-270-437D-5 (1-1708) x SGI_BOVIN (1-646)

OY 789 COAGCAGGTTACTGGGCTGCTC 812
Db 3 ProAla1aLeuLeuGlyLeuLeu 10

RESULT 12
BAC1_MOUSE
ID BAC1_MOUSE STANDARD; PRT; 739 AA.
AC P97302.
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Transcription regulator protein BACH1 (BTB and CNC homolog 1).
GN BACH1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RX MEDLINE=97042438; PubMed=8887638;
RA Oyake T., Itch K., Notohashi H., Hayashi N., Hoshino H., Nishizawa M.,
RA Yamamoto M., Igarashi K.;
RT "Bach proteins belong to a novel family of BTB-basic leucine zipper
RT transcription factors that interact with MafK and regulate
RT transcription through the NF-E2 site."
RL Mol. Cell. Biol. 16:6083-6095(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NMRI; TISSUE=Mammary gland;
RX MEDLINE=2238257; PubMed=12477532;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grumwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S., Krzywinski M.I., Skalska U., Smalilus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Maria M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences".
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Transcriptional regulator that acts as repressor or
CC activator. Binds, in-vitro, to NF-E2 binding sites. Play important
CC roles in coordinating transcription activation and repression by
CC MAFK.

```

CC -!- SUBUNIT: Heterodimer of BACH1 and MAFK.  
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -!- TISSUE SPECIFICITY: Ubiquitous.  
 CC -!- SIMILARITY: Belongs to the bZIP family. CNC subfamily.  
 CC -!- SIMILARITY: Contains 1 BTB/POZ domain.  
 CC  
 CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; D86603; BAA13137.1; -.  
 CC EMBL; BC057894; AAH57894.1; -.  
 CC HSSP; P34707; LSKN.  
 CC TRANSFAC; TC4793; -.  
 CC MGD; MGI:894680; Bach1.  
 CC GO; GO:0005634; C:nucleus; IDA.  
 CC GO; GO:0005515; F:protein binding; IPI.  
 CC GO; GO:0003700; F:transcription factor activity; IDA.  
 CC GO; GO:0006355; P:regulation of transcription, DNA-dependent; IDA.  
 CC InterPro; IPR000210; BTB\_POZ.  
 CC InterPro; IPR008917; Euk\_transcr\_DNA.  
 CC InterPro; IPR002112; LeuZip\_Jun.  
 CC InterPro; IPR004827; TF\_bZIP.  
 CC Pfam; PF00651; BTB; 1.  
 CC Pfam; PF00170; bZIP; 1.  
 CC PRINTS; PR00043; LEUZIPPRJUN.  
 CC SMART; SM00336; BRLZ; 1.  
 CC SMART; SM00225; BTB; 1.  
 CC PROSITE; PS50097; BTB; 1.  
 CC PROSITE; PS0217; bZIP; 1.  
 CC PROSITE; PS00036; bZIP\_BASIC; 1.  
 CC Transcription regulation; Activator; Repressor; DNA-binding;  
 KW Nuclear protein.  
 FT DOMAIN 34 100 BTB.  
 FT DNA\_BIND 565 580 BASIC MOTIF.  
 FT DOMAIN 588 610 LEUCINE-ZIPPER.  
 SQ SEQUENCE 739 AA; 81373 MW; CE2DE60B05F6E32 CRC64;

Alignment Scores:  
 Pred. No.: 30.2 Length: 739  
 Score: 8.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.44% Indels: 0  
 DB: 1 Gaps: 0

US-09-270-437D-5 (1-1708) x BAC1\_MOUSE (1-739)  
 OY 987 ATACGAACCTTGGAGTCAGGTGTT 964  
 Db 239 IleArgThrLeuGlySerGlyVal 246

RESULT 13  
 GYRB\_MYXXA  
 ID GYRB\_MYXXA STANDARD; PRT; 815 AA.  
 AC Q33367.  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE DNA gyrase subunit B (EC 5.99.1.3).  
 GN GYRB.  
 OS Myxococcus xanthus.  
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;  
 OC Cystobacterineae; Myxococcaceae; Myxococcus.  
 OX NCBI\_TaxID=34;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ER-15;  
 RX MEDLINE=98304088; PubMed=9639935;

RA Paitan Y., Boulton N., Ron E.Z., Rosenberg E., Orr E.;  
 RT "Molecular analysis of the DNA gyrase gene from *Mycococcus xanthus*.";  
 RL Microbiology 144:1641-1647(1998).  
 CC -!- FUNCTION: DNA gyrase negatively supercoils closed circular double-  
 CC stranded DNA in an ATP-dependent manner and also catalyzes the  
 CC interconversion of other topological isomers of double-stranded  
 CC DNA rings, including catenanes and knotted rings.  
 CC -!- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining  
 CC of double-stranded DNA.  
 CC -!- SUBUNIT: Made up of two chains. The A chain is responsible for DNA  
 CC breakage and rejoining; the B chain catalyzes ATP hydrolysis. The  
 CC enzyme forms an A2B2 tetramer.  
 CC -!- SIMILARITY: Belongs to the type II topoisomerase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AJ000543; CA04176.1; -.  
 DR HSSP; P06982; 1AJ6.  
 DR InterPro; IPR003594; ATPbind\_ATPase.  
 DR InterPro; IPR002288; DNA\_gyraseB\_C.  
 DR InterPro; IPR000565; DNA\_gyrB.  
 DR InterPro; IPR001241; DNA\_topoisomII.  
 DR InterPro; IPR006171; Toprim dom.  
 DR Pfam; PF00204; DNA\_gyraseB; 1.  
 DR Pfam; PF00986; DNA\_gyraseB; 1.  
 DR Pfam; PF02518; DNA\_gyraseC; 1.  
 DR Pfam; PF01751; HATPase\_C; 1.  
 DR Pfam; PF01751; Toprim; 1.  
 DR PRINTS; PR00418; TP12FAMILY.  
 DR ProDom; PD149633; DNA\_gyraseB\_C; 1.  
 DR SMART; SM00387; HATPase\_C; 1.  
 DR SMART; SM00433; TOP2c; 1.  
 DR TIGRFAMs; TIGR01059; gyrB; 1.  
 DR PROSITE; PS00177; TOPOISOMERASE II; 1.  
 DR Topoisomerase; isomerase; ATP-binding.  
 KW Topoisomerase; isomerase; ATP-binding.  
 SQ SEQUENCE 815 AA; 89636 MW; 3862685FBB805B32 CRC64;

Alignment Scores:  
 Pred. No.: 29.9 Length: 815  
 Score: 8.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.43% Indels: 0  
 DB: 1 Gaps: 0

US-09-270-437D-5 (1-1708) x GYRB\_MYXXA (1-815)

QY 1653 GAGGTGGATCACCTCAGTGGG 1676  
 |||||  
 Db 289 GluGlyGlySerHisLeuSerGly 296

RESULT 14

Y018\_MYCPN  
 ID Y018\_MYCPN STANDARD; PRT; 1030 AA.  
 AC P75093;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical helicase MG018/MG017/MG016 homolog (D12\_orf1030).  
 GN MF020 OR MF134.  
 OS Mycoplasma pneumoniae.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2104;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 29342 / M129;  
 EX MEDLINE=97105885; PubMed=8948633;  
 RA Himmelreich R., Hilbert H., Piagens H., Pirkl E., Li B.-C.,

RA Herrmann R.;  
 RT "Complete sequence analysis of the genome of the bacterium *Mycoplasma pneumoniae*.";  
 RL Nucleic Acids Res. 24:4420-4449(1996).  
 CC -!- SIMILARITY: Belongs to the SNF2/RAD54 helicase family.  
 CC -!- SIMILARITY: TO M.GENITALIUM MG016, MG017 AND MG018.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AE000015; AAB95782.1; -.  
 DR PIR; S73460; S73460.  
 DR InterPro; IPR001410; DEAD.  
 DR InterPro; IPR001650; Helicase\_C.  
 DR InterPro; IPR000330; SNF2\_N.  
 DR InterPro; IPR007527; Znf\_SWIM.  
 DR Pfam; PF00271; helicase\_C; 1.  
 DR Pfam; PF00176; SNF2\_N; 1.  
 DR Pfam; PF04434; SWIM; 1.  
 DR SMART; SM00487; DEXDC; 1.  
 DR SMART; SM00490; HELIC; 1.  
 DR KW Hypothetical protein; Helicase; ATP-binding; Complete proteome.  
 FT NP\_BIND 603 610 ATP (POTENTIAL).  
 FT SITE 702 705 DEAQ BOX.  
 SQ SEQUENCE 1030 AA; 119601 MW; 7272E2B162AF1737 CRC64;

Alignment Scores:  
 Pred. No.: 29 Length: 1030  
 Score: 8.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.43% Indels: 0  
 DB: 1 Gaps: 0

US-09-270-437D-5 (1-1708) x Y018\_MYCPN (1-1030)

QY 1172 AATTGACGGCAGCTGAGTGGA 1195  
 |||||  
 Db 945 AsnLeuThrAlaAlaGluValVal 952

RESULT 15

TOP2\_ARATH  
 ID TOP2\_ARATH STANDARD; PRT; 1473 AA.  
 AC P30182; Q38807;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE DNA topoisomerase II (EC 5.99.1.3).  
 GN TOP2 OR AT3G23890 OR F1013.7.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=95148754; PubMed=7846176;  
 RA Xie S., Lam E.;  
 RT "Characterization of a DNA Topoisomerase II cDNA from Arabidopsis  
 thaliana.";  
 RL Plant Physiol. 106:1701-1702(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia, and cv. Kas-1;  
 RX MEDLINE=95140639; PubMed=7838729;  
 RA Xie S., Lam E.;  
 RT "Abundance of nuclear DNA topoisomerase II is correlated with

RT proliferation in Arabidopsis thaliana.";

RL Nucleic Acids Res. 22:5729-5736(1994).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Columbia;

RX MEDLINE=20363099; PubMed=10907853;

RA Kaneko T., Katoh T., Sato S., Nakamura A., Asamizu E., Tabata S.;

RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.

RT Sequence features of the 4,251,695 bp regions covered by 90 P1, TAC

RT and BAC clones.";

RL DNA Res. 7:217-221(2000).

RN [4]

RP SEQUENCE OF 751-838 FROM N.A.

RA Gerhold D., Parsons A., Hadwiger L.A.;

RL Submitted (XX-1992) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: Control of topological states of DNA by transient

CC breakage and subsequent rejoining of DNA strands. Topoisomerase II

CC makes double-strand breaks.

CC -!- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining

CC of double-stranded DNA.

CC -!- SUBUNIT: Homodimer.

CC -!- MISCELLANEOUS: Eukaryotic topoisomerase I and II can relax both

CC negative and positive supercoils, whereas prokaryotic enzymes

CC relax only negative supercoils.

CC -!- SIMILARITY: Belongs to the type II topoisomerase family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

QY 538 CTTCTTCAGGTTCCGGTCCTTCCTT 515

Db 561 LeuLeuGlnValProSerPheLeu 568

Search completed: July 13, 2004, 12:31:31

Job time : 31.5 secs

EMBL; L21015; AAA65448.1; -;

EMBL; M84854; AAA32877.1; -;

EMBL; AP001297; BAB03006.1; -;

EMBL; U12284; AAC48999.1; -;

EMBL; U12285; AAC49000.1; -;

PIR; S53598; S53599.

HSP; P06786; 1BCW.

InterPro; IPR003594; ATPbind\_ATPase.

InterPro; IPR003957; CBFA\_NFYB\_topis.

InterPro; IPR001241; DNA\_topoisoi.

InterPro; IPR002205; DNA\_topoisoi.

Pfam; PF00204; DNA\_gyraseB; 1.

Pfam; PF00521; DNA\_topoisoi; 1.

Pfam; PF02518; HATase\_C; 1.

PRINTS; PR00615; CCAATSUBUNTA.

PRINTS; PR00418; TP12FAMILY.

ProDom; PD000742; DNA\_topoisoi; 1.

SMART; SM00387; HATase\_C; 1.

SMART; SM00433; TOP2c; 1.

SMART; SM00434; TOP4c; 1.

PROSITE; PS00177; TOPOISOMERASE II; 1.

Isomerase; Topoisomerase; DNA-binding; ATP-binding.

NP\_BIND 161 166

ACT\_SITE 794 794

VARIANT 1213 1213

VARIANT 1245 1245

VARIANT 1299 1299

SEQUENCE 1473 AA; 164106 MW; 00B6C4836E381403 CRC64;

Alignment Scores:

| Pred. No.:             | 27.7    | Length:       | 1473 |
|------------------------|---------|---------------|------|
| Score:                 | 8.00    | Matches:      | 8    |
| Percent Similarity:    | 100.00% | Conservative: | 0    |
| Best Local Similarity: | 100.00% | Mismatches:   | 0    |
| Query Match:           | 1.44%   | Indels:       | 0    |
| DB:                    | 1       | Gaps:         | 0    |



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: July 13, 2004, 12:12:42 ; Search time 87 Seconds  
(without alignments)  
12247.273 Million cell updates/sec

Title: US-09-270-437D-5  
Perfect score: 3110  
Sequence: 1 aggcagctgcgcacgcgc.....attcttcaggttttaaaa 1708

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1279676 segs, 311918243 residues

Total number of hits satisfying chosen parameters: 2559352

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=frame+n2p.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.spool\_p/US09270437/runat\_13072004\_121924\_9724/app\_query.fasta\_1.1863  
-DB=Published Applications AA -QFMT=fastan -SUFFIX=rapp -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=-1 -MATRIX=blosu62  
-TRANS=human40 cdi -LISN=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTPMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09270437@cgn\_13072004\_13072004\_121924\_9724  
-NCPU=6 -ICPU=3 -NO WMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA.\*

- 1: /cgn2\_6/ptodata/1/pubppa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubppa/FCI\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubppa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubppa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubppa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubppa/US08\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubppa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubppa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubppa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubppa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubppa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubppa/US09A\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubppa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubppa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubppa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubppa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubppa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubppa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
|------------|-------|-------------|--------|----|-------------|

RESULT 1  
US-10-313-986-500  
; Sequence 500, Application US/10313986  
; Publication No. US20030236209A1  
; GENERAL INFORMATION:  
; APPLICANT: Foy, Teresa M.  
; APPLICANT: McNabb, Andria  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Wang, Tongtong  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C19  
; CURRENT APPLICATION NUMBER: US/10/313,986  
; CURRENT FILING DATE: 2002-12-04  
; NUMBER OF SEQ ID NOS: 560  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 500  
; LENGTH: 577  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-313-986-500  
Alignment Scores:

|    |        |      |     |    |                     |                   |
|----|--------|------|-----|----|---------------------|-------------------|
| 1  | 2226   | 71.6 | 577 | 15 | US-10-313-986-500   | Sequence 500, App |
| 2  | 2208   | 71.0 | 577 | 9  | US-09-873-637-2     | Sequence 2, Appli |
| 3  | 1639   | 52.7 | 579 | 9  | US-09-735-705-348   | Sequence 348, App |
| 4  | 1639   | 52.7 | 579 | 9  | US-09-850-716A-348  | Sequence 348, App |
| 5  | 1639   | 52.7 | 579 | 9  | US-09-897-778-348   | Sequence 348, App |
| 6  | 1639   | 52.7 | 579 | 9  | US-09-897-778-446   | Sequence 446, App |
| 7  | 1639   | 52.7 | 579 | 9  | US-09-897-778-449   | Sequence 449, App |
| 8  | 1639   | 52.7 | 579 | 12 | US-10-007-700-348   | Sequence 348, App |
| 9  | 1639   | 52.7 | 579 | 12 | US-10-007-700-446   | Sequence 446, App |
| 10 | 1639   | 52.7 | 579 | 12 | US-10-007-700-449   | Sequence 449, App |
| 11 | 1639   | 52.7 | 579 | 14 | US-10-117-982-348   | Sequence 348, App |
| 12 | 1639   | 52.7 | 579 | 14 | US-10-117-982-446   | Sequence 446, App |
| 13 | 1639   | 52.7 | 579 | 14 | US-10-117-982-449   | Sequence 449, App |
| 14 | 1639   | 52.7 | 579 | 14 | US-10-117-982-480   | Sequence 480, App |
| 15 | 1639   | 52.7 | 579 | 15 | US-10-313-986-446   | Sequence 446, App |
| 16 | 1639   | 52.7 | 579 | 15 | US-10-313-986-449   | Sequence 449, App |
| 17 | 1639   | 52.7 | 579 | 15 | US-10-313-986-480   | Sequence 480, App |
| 18 | 1639   | 52.7 | 579 | 15 | US-10-313-986-486   | Sequence 486, App |
| 19 | 1639   | 52.7 | 586 | 9  | US-09-850-716A-427  | Sequence 427, App |
| 20 | 1639   | 52.7 | 586 | 9  | US-09-897-778-427   | Sequence 427, App |
| 21 | 1639   | 52.7 | 586 | 12 | US-10-007-700-427   | Sequence 427, App |
| 22 | 1639   | 52.7 | 586 | 14 | US-10-117-982-427   | Sequence 427, App |
| 23 | 1639   | 52.7 | 586 | 15 | US-10-313-986-427   | Sequence 427, App |
| 24 | 1639   | 52.7 | 589 | 15 | US-10-313-986-486   | Sequence 486, App |
| 25 | 1637   | 52.6 | 579 | 9  | US-09-735-705-176   | Sequence 176, App |
| 26 | 1637   | 52.6 | 579 | 9  | US-09-850-716A-176  | Sequence 176, App |
| 27 | 1637   | 52.6 | 579 | 9  | US-09-897-778-176   | Sequence 176, App |
| 28 | 1637   | 52.6 | 579 | 10 | US-09-466-396A-176  | Sequence 176, App |
| 29 | 1637   | 52.6 | 579 | 12 | US-10-007-700-176   | Sequence 176, App |
| 30 | 1637   | 52.6 | 579 | 14 | US-10-117-982-176   | Sequence 176, App |
| 31 | 1637   | 52.6 | 579 | 15 | US-10-313-986-176   | Sequence 176, App |
| 32 | 1635   | 52.6 | 579 | 14 | US-10-117-982-484   | Sequence 484, App |
| 33 | 1635   | 52.6 | 579 | 15 | US-10-313-986-484   | Sequence 484, App |
| 34 | 1560   | 50.2 | 422 | 16 | US-10-408-765A-2088 | Sequence 2088, Ap |
| 35 | 1530.5 | 49.2 | 620 | 9  | US-09-764-864-1116  | Sequence 1116, Ap |
| 36 | 1524.5 | 49.0 | 587 | 15 | US-10-313-986-501   | Sequence 501, App |
| 37 | 1464   | 47.1 | 556 | 14 | US-10-097-340-147   | Sequence 147, App |
| 38 | 1464   | 47.1 | 556 | 16 | US-10-648-593-182   | Sequence 182, App |
| 39 | 1355.5 | 43.6 | 555 | 15 | US-10-262-445-40    | Sequence 40, Appl |
| 40 | 1258   | 40.5 | 261 | 9  | US-09-764-864-1114  | Sequence 1114, Ap |
| 41 | 1228   | 39.5 | 250 | 9  | US-09-764-864-1532  | Sequence 1532, Ap |
| 42 | 653    | 21.0 | 171 | 9  | US-09-764-864-1119  | Sequence 1119, Ap |
| 43 | 510    | 16.4 | 192 | 9  | US-09-764-864-1117  | Sequence 1117, Ap |
| 44 | 502    | 16.1 | 171 | 9  | US-09-764-864-1536  | Sequence 1536, Ap |
| 45 | 361    | 11.6 | 81  | 14 | US-10-117-982-476   | Sequence 476, App |



QY 272 AAGGAGGGGCGCCATCCGCAACATCAAAAAACAGACCCAGTCCAGATAGACGTGCAT 331  
DB 213 LysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHis 232  
QY 332 AGGAGGAGAACGCGAGCTGAGCTGAAAGCCCATCAGTGTGCACTCCACCCCTGAGGGC 391  
DB 233 ArgLysGluAsnAlaGlyAlaGluGlyAlaIleSerValHisSerThrProGluGly 252  
QY 392 TGCCTCCCGCTTGAAGATGATCTGGAGATTATGCAATAAGAGGCTAAGGACACCAAA 451  
DB 253 CysSerSerAlaCysLysMetIleLeuGluIleMetHisLysGluAlaLysAspThrLys 272  
QY 452 ACGCTGACGAGGTTCCCTGAAAGATCTGCGCCCATATAACTTTGTAGGCGCTCATTT 511  
DB 273 ThrAlaAspGluValProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIle 292  
QY 512 GCGAAGGAAGACGCAACTCAAGAAGGTAGACAGATACCGAGACAAAAAATCACCATC 571  
DB 293 GlyLysGluGlyArgAsnLeuLysValGluGlnAspThrGluThrLysIleThrIle 312  
QY 572 TCCTGTTGCAAGACCTTACCTTTACACCTGAGAGGACCATCCTGTGAAGGGGCGCC 631  
DB 313 SerSerLeuGlnAspLeuThrLeuTyAsnProGluArgThrIleThrValLysGlyAla 332  
QY 632 ATCGAGATTCTGCGAGGCGAGCAGGAAATATGAAGAAAGTTGCGGAGGCGCTATGAG 691  
DB 333 IleGluAsnCysArgAlaGluGlnGluIleMetLysLysValArgGluAlaTyGlu 352  
QY 692 AATGATGTGGCTGCATGAGC-----TCTCACTGATCCCTGCGCTGAACCTGGCTGCT 745  
DB 353 AsnAspValAlaAlaMetSerLeuGlnSerHisLeuIleProGlyLeuAsnLeuAlaAla 372  
QY 746 GTAGTCTTTTCCAGCTTCTCCAGCGAGTCCCGCGCGCTCCAGCAGGCTTACTGGG 805  
DB 373 ValGlyLeuPheProAlaSerSerSerAlaValProProProProSerSerValThrGly 392  
QY 806 GCTGCTCCTTAGTCTCTTTATGAGGCTCCCGAGCAGGAGATGCTGCGAGTCTTTATC 865  
DB 393 AlaAlaProTySerSerPheMetGlnAlaProGluGlnGluMetValGlnValPheIle 412  
QY 866 CCGGCGCCAGGAGTGGGCGCATCATCGGCAAGAGGGGCGAGCACATCAACACGCTCTCC 925  
DB 413 ProAlaGlnAlaValGlyAlaIleIleGlyLysLysGlyGlnHisIleLysGlnLeuSer 432  
QY 926 CGGTTTGCAGCGCTCATCAAGATTGCACACCGCAACACCTGACTCCCAAAAGTTCGT 985  
DB 433 ArgPheAlaSerAlaSerIleLysIleAlaProProGluThrProAspSerLysValArg 452  
QY 986 ATGGTTATCATCATCGGACCGCCAGAGGCCCAATTCAAAGGCTCAGGGAAGATCTATGC 1045  
DB 453 MetValIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyGly 472  
QY 1046 AAACCTCAAGGAGGAGAACTTCTTGTCTCCAGGAGGAGTGAAGCTGGAGACCCACATA 1105  
DB 473 LysLeuLysGluGluAsnPheGlyProLysGluGluValLysLeuGluThrHisIle 492  
QY 1106 CGTGTGCCAGCATCAGCAGCTGGCGGGTCAATTGGCAAAAGGTGGAATAACCGTGAACGAG 1165  
DB 493 ArgValProAlaSerAlaAlaGlyArgValIleGlyLysGlyLysThrValAsnGlu 512  
QY 1166 TTGCAAGATTTCACGCGAGCTGAGGTGTAGTACCAAGACGACGACCCCTGATGAGAAC 1225  
DB 513 LeuGlnAsnLeuThrAlaAlaGluValValProArgAspGlnThrProAspGluAsn 532  
QY 1226 GACCAAGTTCATCGTGAATAATCATCGGACATTTCTATGCCAGTCAGATGGCTCAACGGAG 1285  
DB 533 AspGlnValIleValLysIleIleGlyHisPheTyAlaSerGlnMetAlaGlnArgLys 552  
QY 1286 ATCCGAGACATCTCGGCCAGGTTAAGCAGCAGCATCAGAAGGCGCAGAGTAACCGAGCC 1345  
DB 553 IleArgAspIleLeuAlaGlnValLysGlnGlnHisGlnLysGlyGlnSerAsnLeuAla 572  
QY 1346 CAGGACGAGGAGAG 1360

DB 573 GlnAlaArgArgLys 577  
RESULT 3  
US-09-735-705-348  
; Sequence 348, Application US/09735705  
; Patent No. US20020052329A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Ligu  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy  
; APPLICANT: Panger, Gary R.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Panger, Neil  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.455C14  
; CURRENT APPLICATION NUMBER: US/09/735,705  
; CURRENT FILING DATE: 2000-12-12  
; NUMBER OF SEQ ID NOS: 419  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 348  
; LENGTH: 579  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-735-705-348  
Alignment Scores:  
Pred. No.: 1,266-121 Length: 579  
Score: 1639.00 Matches: 336  
Percent Similarity: 83.89% Conservative: 44  
Best Local Similarity: 74.17% Mismatches: 57  
Query Match: 52.70% Indels: 16  
Gaps: 9  
DB:  
US-09-270-437D-5 (1-1708) x US-09-735-705-348 (1-579)  
QY 32 CCGGAGGACCATCATGAAGCTGAATGCGCACCATGTCGAGAACCATCCCTGAAGGTCTCC 91  
DB 133 ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla 152  
QY 92 TACATCCCGATCAGCAGATAGCA---CAGGGACCTCAGATGGCGCGCGAGG----- 142  
DB 153 TyrlleProAspGluThrAlaAlaGlnGlnAsnProLeuGlnProArgGlyArgArg 172  
QY 143 GGCTTTGGCTCTCGGGTTCAGCCCGCCAGGGCTCACCTGTGTCAGCGGGGGCCGAGCC 202  
DB 173 GlyLeuGlyGlnArgGlySerSerArgGlnGlySerPro-----GlySerValSer 189  
QY 203 AAGCAGCAGCAAGTGACATCCCTTCCTGCTCGTCCACCCAGTATGTTGGTGC 262  
DB 190 LysGlnLysProCysAspLeuProLeuArgLeuValProThrGlnPheValGlyAla 209  
QY 263 ATTATTGGCAAGGAGGGGCGCCACCATCCGCAACATCACAAACAGACCCAGTCCAAAGATA 322  
DB 210 IlelleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIle 229  
QY 323 GACGTGCATAGGAGGAGAACCGAGGTGCGAGCTGAGTGAATAAGCCATCATGTGCTCACC 382  
DB 230 AspValHisArgLysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThr 249  
QY 383 CCTGAGGCTGCTCCCTCCGCTTGTAGATGATCTTGAGATTATGCATAAAGAGGCTTAAG 442  
DB 250 ProGluGlyThrSerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGln 269  
QY 443 GACACCAAAACGGCTGACGAGGTTCCCTCAAGATCTTCGCCCATATAAATCTTGTAGGG 502



```

Db 390 ProSerAlaMetThr-----ProProTyrProGlnPheGluGlnSer--GluThrGlu 406
QY 848 ATGTGTCAGAGGTTATCCCGCCAGCAGTGGCGCCATCATCGGCAAGAGGGGCAG 907
Db 407 ThrValHisLeuPheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGln 426
QY 908 CACATCAAAACAGCTCTCCGGTTTGCACGCGCCCTCCATCAAGATTGCACCAACCGAAACA 967
Db 427 HisLeuLysGlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAla 446
QY 968 CCTGACTCAAGTTCGTATGTTATCATCTCGACCGCCAGAGGCCCAATTCAGGCT 1027
Db 447 ProAspAlaLysValArgMetValIleIleThrGlyProGlnAlaGlnPheLysAla 466
QY 1028 CAGGGAAGAATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTG 1087
Db 467 GlnGlyArgIleTyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluVal 486
QY 1088 AAGCTGGAGACCCATACGTGTGCAGCATCAGAGCTGCCCGGGTCATGGCAAGGT 1147
Db 487 LysLeuGluAlaHisIleArgValProSerPheAlaGlyArgValIleGlyLysGly 506
QY 1148 GGAAAAACGGTGAACGAGTTGCAGAAATTTGACGGCAGCTGAGGTGGTAGTACCAAGAGAC 1207
Db 507 GlyLysThrValAsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAsp 526
QY 1208 CAGACCCCTGATGAAACGACAGCATCTCTGTAATATCATCGACATTCATCGCCAGT 1267
Db 527 GlnThrProAspGluAsnAspGlnValValValLysIleThrGlyHisPheTyrAlaCys 546
QY 1268 CAGATGGCTCAACGAAAGTCCGACATCTCTGGCCAGGTTAAGCAG--CAGCATCAG 1324
Db 547 GlnValAlaGlnArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGln 566
QY 1325 AAGGGA---CAGAGTAAACAGGCCCGCAGGCGAGGGAAG 1360
Db 567 LysAlaLeuGlnSerGlyProGlnSerArgArgLys 579

RESULT 5
US-09-897-778-348
; Sequence 348, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Warnerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C16
; CURRENT APPLICATION NUMBER: US/09/897,778
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 348
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-897-778-348

Alignment Scores:
Pred. No.: 1,26e-121 Length: 579
Score: 1639.00 Matches: 336
Percent Similarity: 83.89% Conservative: 44
Best Local Similarity: 74.17% Mismatches: 57
Query Match: 52.70% Indels: 16
DB: 9 Gaps: 9
```

```

US-09-270-437D-5 (1-1708) x US-09-897-778-348 (1-579)
QY 32 CGGGGAGCCATCATGAAGCTGAATGGCCACCACTGTTGGAGAACATGCTCCCTGAAGTCTCC 91
Db 133 ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla 152
QY 92 TACATCCCCGATGAGCAGATAGCA---CAGGGACCTGAGAAATGGGGCCGAGGG----- 142
Db 153 TyrIleProAspGluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArg 172
QY 143 GCGTTTGGTCTCTCGGGGTTCAGCCCGCCAGGCTCACCTGTGGCAGCGGGGCCCGCCAGCC 202
Db 173 GlyLeuGlyGlnArgGlySerSerArgGlnGlySerPro-----GlySerValSer 189
QY 203 AAGCAGCAGCAAGTGGACATCCCCCTTCGGCTCTGGTGTCCCAACCCAGTATGGGTGCC 262
Db 190 LysGlnLysProCysAspLeuProLeuArgLeuValProThrGlnPheValGlyAla 209
QY 263 ATTATTGGCAAGAGGGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCCCAAGATA 322
Db 210 IleileGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIle 229
QY 323 GAGCTGCATAGGAAGAGAACCGCAGGTGCAGCTGAAAGAACCCATCATGTGTGCCTCCACC 382
Db 230 AspValHisArgLysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThr 249
QY 383 CCTGAGGGTGTCTCCCTCGCTGTAGATGATCTTGGAGATTATGCAATAAGAGGCTAAG 442
Db 250 ProGluGlyThrSerAlaAlaCysLysSerIleLeuGluIleMethHisLysGluAlaGln 269
QY 443 GACACAAAACGGCTGACGAGGTTCCTCTGAAGATCCTGCGCCCATATAATTAATCTGTAGGG 502
Db 270 AspIleLysPheThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGly 299
QY 503 CGTCTCATTTGGCAAGAGACGGAACCTGAGAGGTAGAGCAGATACCGAGACAAAA 562
Db 290 ArgLeuIleGlyLysGluGlyArgAsnLeuLysLysIleGluGlnAspThrAspThrLys 309
QY 563 ATCACCATCTCTCTGTTGCAAGACCTTACCTTTACACCCCTGAGAGGACCATCACTGTG 622
Db 310 IleThrIleSerProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrVal 329
QY 623 AAGGGGGCCATCGAGAATTGTTTCAGGGCCGAGCAGAGAAATATAGAAAGTTCGGGAG 682
Db 330 LysGlyAsnValGluThrCysAlaLysAlaGluGluGluIleMethLysIleAsglu 349
QY 683 GCCTATGAGAAATGATGTGGTGCATGAGC-----TCTCACCTGATCCCTGCGCTCAAC 736
Db 350 SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn 369
QY 737 CTGGCTGCTGTAGGTCTTTTCCAGGCTTCATCCAGCGCAGTCCCGCCG-----CCT 787
Db 370 LeuAsnAlaLeuGlyLeuPheProProThrSerGlyMetProProThrSerGlyPro 389
QY 788 CCCAGCAGCGTTACTGGGGCTGTCTCCCTATAGCTCTCTTATGAGGCTCCCGAGCAGGAG 847
Db 390 ProSerAlaMetThr-----ProProTyrProGlnPheGluGlnSer---GluThrGlu 406
QY 848 ATGTGTCAGGTGTTTATCCCGCCAGGCGAGTGGCGCCATCATCGGCAAGAGGGGCAG 907
Db 407 ThrValHisLeuPheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGln 426
QY 908 CACATCAAAACAGCTCTCCCGTTTGGCAGCGCTCCATCAAGATTGCACCAACCGCAACA 967
Db 427 HisIleLysGlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAla 446
QY 968 CCTGACTCCAAAGTTCGTATGTTATCATCTCGACCGCCAGAGGCCCAATTCAGGCT 1027
Db 447 ProAspAlaLysValArgMetValIleIleThrGlyProGlnAlaGlnPheLysAla 466
QY 1028 CAGGGAAGAATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTG 1087
```

```

Db      467 GlnGlyArgIleTyGlyLysIleLysGluGluAenPheValSerProLysGluGluVal 486
QY      1088 AAGCTGGAGACCCACATACGTGTGCCACATCAGCAGCTGGCCGGTCAATTGGCAAGGT 1147
Db      487 LysLeuGluAlaHisIleArgValProSerPheAlaAlaGlyArgValIleGlyGlyGly 506
QY      1148 GAAAAACGGTGAACAGTTCAGAAATTTGACGGCAGCTGAGTGGTGTAGTACCAGAGAC 1207
Db      507 GlyLysThrValAsnGluLeuGlnAenLeuSerAlaGluValValProArgAsp 526
QY      1208 CAGACCCCTGTAGAGAACACAGCTCATCGTAAATCATCGGACATTTCTATGCCAGT 1267
Db      527 GlnThrProAspGluAsnAspGlnValValValLysIleThrGlyHisPheTyAlaCys 546
QY      1268 CAGATGGCTAACGGAGATCCGAGACATCTCTGGCCCGCCAGGTTAAGCAG---CAGCATCAG 1324
Db      547 GlnValAlaGlnArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGln 566
QY      1325 AAGGGA---CAGAGTAACACAGGCCCGCAGCAGCGAGGAG 1360
Db      567 LysAlaLeuGlnSerGlyProProGlnSerArgArgLys 579

RESULT 6
US-09-897-778-446
; Sequence 446, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Derrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C16
; CURRENT APPLICATION NUMBER: US/09/897,778
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 446
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-897-778-446

Alignment Scores:
Pred. No.: 1,26e-121 Length: 579
Score: 1639.00 Matches: 336
Percent Similarity: 83.89% Conservative: 44
Best Local Similarity: 74.17% Mismatches: 57
Query Match: 52.70% Indels: 16
DB: 9 Gaps: 9

US-09-270-437D-5 (1-1708) x US-09-897-778-446 (1-579)
QY      32 CGGGAGCCATCATGAAGCTGAATGGCCACAGTTGGAGAACCATGCCCTGAAGGTCTCC 91
Db      133 ArgGlnAlaLeuApyLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla 152
QY      92 TACATCCCGATGACGATAGCA---CAGGGACCTGAGAAATGGCGCCGAGG-----142
Db      153 TyrIleProAspGluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArg 172
QY      143 GGCTTTGCTCTCGGGGTGACCCCGCCAGGGCTCACCTGTGCAGCGGGGGCCCGAGCC 202
Db      173 GlyLeuGlnArgGlySerSerArgGlnGlySerPro-----GlySerValSer 189
QY      203 AAGCAGCAGCAAGTGGACATCCCGCTTCGGCTCGCTGCGCCACCCAGCATGTGGGTGCC 262

```

```

Db      190 LysGlnLysProCysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAla 209
QY      263 ATTATTGGCAAGGAGGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCCAGATA 322
Db      210 IleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIle 229
QY      323 GACGTCCATAGGAAGAGAACGAGGTGCAGTGAAGAAAGCCATCAGTGTGACATCCACCC 382
Db      230 AspValHisArgLysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThr 249
QY      383 CTGAGGGGTGCTCCCTCCCTTTGTAAGATGATCTTGGAGATTATGCATAAAGAGCTAAG 442
Db      250 ProGluGlyThrSerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGln 269
QY      443 GACACCAAAACCGCTGACGAGGTTCCTCTGAAGATCTCGGCCCATATAAATCTTTGAGGG 502
Db      270 AspIleLysPheThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGly 289
QY      503 CGTCTCATTTGGCAAGAGAGCGGAACCTGAAGAGAGGTAGAGCAAGATACCGAGACAAA 562
Db      290 ArgLeuIleGlyLysGluGlyArgAsnLeuLysLysIleGluGlnAspThrAspThrLys 309
QY      563 ATCACCATCTCTCGTTGCAAGACCTTACCTTTTACAACCTGAGAGAGACCATCACTCG 622
Db      310 IleThrIleSerProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrVal 329
QY      623 AAGGGGCCATCGAGAATTGTGAGGGCCGAGGAAATAATGAAGAAATTCGGGAG 682
Db      330 LysGlyAsnValGluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGlu 349
QY      683 GCCTATGAGAAATGATGTGCTGCCATGAGC-----TCTCACCTGATCCCTGGCTGAC 736
Db      350 SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn 369
QY      737 CTGGCTGTGTAGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCCG-----CCT 787
Db      370 LeuAsnAlaLeuGlyLeuPheProProThrSerGlyMetProProProThrSerGlyPro 389
QY      788 CCACAGCGGTACTTGGGGTCTCCCTATAGTCTCTTATGTCAGGCTCCGAGCAGAG 847
Db      390 ProSerAlaMetThr-----ProProTyrProGlnPheGluGlnSer---GluThrGlu 406
QY      848 ATGGTGCAGGTCTTTATCCCGCCAGCAGTGGCGCCATCATCGGCAAGAGGGGCGAG 907
Db      407 ThrValHisLeuPheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGln 426
QY      908 CACATCAAAACAGCTCTCCGGTTGCGAGCGCTCCATCAAGATTCACACCCGAAACA 967
Db      427 HisIleLysGlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAla 446
QY      968 CTGACTCCAAAGTTCGTATGGTATCATCATCTGACCGCCGAGAGGCCCAATTCAGGCT 1027
Db      447 ProAspAlaLysValArgMetValIleIleThrGlyProGluAlaGlnPheLysAla 466
QY      1028 CAGGGAAGAATCTATGGCAACTCAAGGAGAGAACTTCTTTGGTCCCAAGAGAGAGTG 1087
Db      467 GlnGlyArgIleTyrGlyLysIleLysGluLysLeuAsnPheValSerProLysGluGluVal 486
QY      1088 AAGCTGGAGACCCACATACGTGTGCAGCATCAGCAGCTGGCCGGTCAATTCGCAAGGT 1147
Db      487 LysLeuGluAlaHisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGly 506
QY      1148 GAAAAACGGTGAACAGTTCAGAAATTTGACGGCAGCTGAGTGGTAGTACCAAGAGAC 1207
Db      507 GlyLysThrValAsnGluLeuGlnAsnLeuSerAlaGluValValProArgAsp 526
QY      1208 CAGACCCCTGTAGAGAACGAGCAGGTCTATCGTGAATATCATCGACATTTCTATGCCAGT 1267
Db      527 GlnThrProAspGluAsnAspGlnValValLysIleThrGlyHisPheTyAlaCys 546
QY      1268 CAGATGGCTAACGAGAGATCCGAGACATCTCGGCCCGCAGGTTAAGCAG---CAGCATCAG 1324
Db      547 GlnValAlaGlnArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGln 566

```

QY 1325 AAGGGA---CAGAGTAACCGCCAGCCAGCCAGCGAGGAAG 1360  
 Db |||||  
 567 LysAlaLeuGlnSerGlyProProGlnSerArgArgLys 579

## RESULT 7

US-09-897-778-449  
 ; Sequence 449, Application US/0989778  
 ; Patent No. US20020147143A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Tongtong  
 ; APPLICANT: Marnerakis, Margarita  
 ; APPLICANT: Fanger, Gary R.  
 ; APPLICANT: Vedvick, Thomas S.  
 ; APPLICANT: Carter, Barrick  
 ; APPLICANT: Watanabe, Yoshihiro  
 ; APPLICANT: Henderson, Robert A.  
 ; APPLICANT: Peckham, David W.  
 ; APPLICANT: Fanger, Neil  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
 ; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
 ; FILE REFERENCE: 210121.455C16  
 ; CURRENT APPLICATION NUMBER: US/09/897,778  
 ; CURRENT FILING DATE: 2001-06-28  
 ; NUMBER OF SEQ ID NOS: 467  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 449  
 ; LENGTH: 579  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-897-778-449

Alignment Scores:  
 Pred. No.: 1,266-121 Length: 579  
 Score: 1639.00 Matches: 336  
 Percent Similarity: 83.89% Conservative: 44  
 Best Local Similarity: 74.17% Mismatches: 57  
 Query Match: 52.70% Indels: 16  
 DB: 9 Gaps: 9  
 US-09-270-437D-5 (1-1708) x US-09-897-778-449 (1-579)

QY 32 CGGGAGCCATCAAGCTGAATGCCACCGCAGTGGAGAACCATGCCCTGAAGTCTCC 91  
 Db |||||  
 133 ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla 152  
 QY 92 TACATCCCGCATGACCATAGCA---CAGGACCTGAGATGGCGCGGAGGG----- 142  
 Db |||||  
 153 TyrIleProAspGluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArg 172  
 QY 143 GCCTTTGGCTCTCGGGGTGAGCCCGCCAGGCTCACCTGTGGCAGCGGGGGCCAGCC 202  
 Db |||||  
 173 GlyLeuGlyGlnArgGlySerSerArgGlnGlySerPro-----GlySerValSer 189  
 QY 203 AAGCAGCAGCATGTCGATCCCTTGGCTCTGGTGGCCACCCAGTATGTGGTGCC 262  
 Db |||||  
 190 LysGlnLysProCysAspLeuProLeuArgLeuValProThrGlnPheValGlyAla 209  
 QY 263 ATTATTGGCAAGAGGGGCCACCATCCGCAACATCACAAACAGACCCAGTCCAAAGATA 322  
 Db |||||  
 210 IleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIle 229  
 QY 323 GACGTGCATPAGAGGAGAACCCAGGTGCAGCTGAAAAGCCATCAGTGTGCATCCACC 382  
 Db |||||  
 230 AspValHisArgLysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThr 249  
 QY 383 CCTGAGGGTGTCTCCCTGTAGATGATCTTGGAGATTATGCATAAGAGGCTAAG 442  
 Db |||||  
 250 ProGluGlyThrSerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGln 269  
 QY 443 GACACCAAAACCGGTGACGAGGTCCCTTGAGATCCCTGGCCCATTAACCTTTGTAGGG 502  
 Db |||||  
 270 AspIleLysPheThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGly 289

QY 503 CGTCTCATTGGCAAGGACGACGGAACCTGAAGAGGTAGAGACATCCGAGACAAA 562  
 Db |||||  
 290 ArgLeuIleGlyLysGluGlyArgAsnLeuLysLysIleGluAsnThrAspThrLys 309  
 QY 563 ATCACCATTCTCTCGTTGCAAGACCTTACCCCTTTCAACCCCTGAGAGGACCATCTGTG 622  
 Db |||||  
 310 IleThrIleSerProLeuGlnGluLeuThrLeuLysAsnProGluArgThrIleThrVal 329  
 QY 623 AAGGGGCCATCGAAGTGTTCAGGGCCGACGAGGAAATATGAAGAAGTTCGGGAG 682  
 Db |||||  
 330 LysGlyAsnValGluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGlu 349  
 QY 683 GCCTATGAGATCATCTGGTCCCATGAC-----TCTCAGCTGATCCCTGCGCTGAAC 736  
 Db |||||  
 350 SerTyrGluAsnAspIleAlaSerMetAsnLeuAlaHisLeuIleProGlyLeuAsn 369  
 QY 737 CTGGCTGCTAGTGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCG-----CCT 787  
 Db |||||  
 370 LeuAsnAlaLeuGlyLeuPheProProThrSerGlyMetProProProThrSerGlyPro 389  
 QY 788 CCCAGCAGGTTACTGGGCTGCTCCCTATAGTCTCTTTATGAGGCTCCGAGCAGGAG 847  
 Db |||||  
 390 ProSerAlaMetThr-----ProProTyrProGlnPheGluGlnSer---GluThrGlu 406  
 QY 848 ATGCTGCAGGTGTTTATCCCGCCAGGAGTGGCGCCATCATCGCCCAAGAGGGGCGAG 907  
 Db |||||  
 407 ThrValHisLeuPheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGln 426  
 QY 908 CACATCAAAAGCTCTCCCGTTTGCAGCGCTCCATCAAGATTGCACACCCGAAACA 967  
 Db |||||  
 427 HisIleLysGlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAla 446  
 QY 968 CTTGACTCCAAAGTTCGTATGTTATCATCATCTGGACCGCCAGAGGCCCAATTCAGGCT 1027  
 Db |||||  
 447 ProAspAlaLysValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAla 466  
 QY 1028 CAGGGAAGATCTATCGCAAACTCAAGGAGGAGAACTCTTTGGTCCCAAGGAGGAAGTG 1087  
 Db |||||  
 467 GlnGlyArgIleTyrGlyLysIleLysGluLysPheValSerProLysGluGluVal 486  
 QY 1088 AAGCTGAGACCCACATAGTGTGCCAGCATCAGCAGCTGGCGCGCTCATTCGCAAGGT 1147  
 Db |||||  
 487 LysLeuGluAlaHisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGly 506  
 QY 1148 GGAACCAAGCTGAACAGTTCAGAAATTTGACGCGCAGCTGAGGTGGTAGTACCAAGAGAC 1207  
 Db |||||  
 507 GlyLysThrValAsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAsp 526  
 QY 1208 CAGACCCCTGATCAGAACCGACCGAGTTCGTGAAATCATCGACATTTCTATGCGAGT 1267  
 Db |||||  
 527 GlnThrProAspGluAsnAspGlnValValLysIleThrGlyHisPheTyrAlaCys 546  
 QY 1268 CAGATGGCTCAACGGGAAGATCCGAGACATCTCGCCCGCAGCTTAAGCAG---CAGCATCAG 1324  
 Db |||||  
 547 GlnValAlaGlnArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGln 566  
 QY 1325 AAGGGA---CAGAGTAACCGAGGCCCGCCAGGACCGAGGAAG 1360  
 Db |||||  
 567 LysAlaLeuGlnSerGlyProProGlnSerArgArgLys 579

## RESULT 8

US-10-007-700-348  
 ; Sequence 348, Application US/10007700  
 ; Publication No. US2003006497A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Tongtong  
 ; APPLICANT: Wang, Aijun  
 ; APPLICANT: Skeiky, Yasir A.W.  
 ; APPLICANT: Li, Samuel X.  
 ; APPLICANT: Kalos, Michael D.  
 ; APPLICANT: Henderson, Robert A.  
 ; APPLICANT: McNeill, Patricia D.



```

; APPLICANT: Fanger, Neil
; APPLICANT: Retter, Marc W.
; APPLICANT: Durham, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darriek
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Peckman, David W.
; APPLICANT: Cai, Feng
; APPLICANT: Foy, Teresa M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C17
; CURRENT APPLICATION NUMBER: US/10/007,700
; CURRENT FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 348
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-007-700-348

Alignment Scores:
Pred. No.: 1,268-121 Length: 579
Score: 1639.00 Matches: 336
Percent Similarity: 83.8% Conservative: 44
Best Local Similarity: 74.17% Mismatches: 57
Query Match: 52.70% Indels: 16
DBs: 12 Gaps: 9

US-09-270-437D-5 (1-1708) x US-10-007-700-348 (1-579)

```

```

QY 32 CGGGAGGACATCAAGTGAATGGCCACCGAGTGGAGAACCATGCCCTCAAGGCTCC 91
DB 133 ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla 152
QY 92 TACATCCCGATGAGCAGATAGCA--CAGGACCTGAGATGGCCCGCAGGG-- 142
DB 153 TyrIleProAspGluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArg 172
QY 143 GCCTTGCTCTCGGGGTGACGCCCGCCAGGGCTCACCTGTGGCAGCGGGGCCCGCCAGCC 202
DB 173 GlyLeuGlyGlnArgGlySerArgGlnGlySerPro-----GlySerValSer 189
QY 203 AAGCAGCAGAGTGGACATCCCTTCGGCTCTGCTGCGCCACCGAGTATGTGGTGCC 262
DB 190 LysGlnLysProCysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAla 209
QY 263 ATTATTGGCAAGGAGGGGCCACCATCGCAACATCAAAAACAGACCCAGTCCAGATA 322
DB 210 IleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIle 229
QY 323 GACGTGCATAGAGGAGAGCAGGTGCGAGTGAAGAACCCATCAGTGTGCTCACC 382
DB 230 AspValHisArgLysGluAsnAlaGlyAlaAlaGlySerIleThrIleLeuSerThr 249
QY 383 CCTGAGGCTCTCTCCGCTGTGAAGATGATCTTGAGATATTGACATAAGAGGCTAAG 442
DB 250 ProGluGlyThrSerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGln 269
QY 443 GACACCAAAACGGCTGACAGGTTCCCTGAAGATCTGCGCCCATATAAATCTTTAGGG 502
DB 270 AspIleLysPheThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGly 289
QY 503 CGTCTCATTTGGCAAGGAGGAGCAACCTGAAGAGGTAGAGCAAGATACCGAGACAAA 562
DB 290 ArgLeuIleGlyLysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLys 309
QY 563 ATCACCATCTCTCTGTCAGACCTTACCTTTACAAACCTCGAGAGGACCATCTGTG 622
DB 310 IleThrIleSerProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrVal 329

```

```

QY 623 AAGGGGCGCATCGAGAAATTGTTGCAGGGCCGAGCAAGAAATAATGAAGAAAGTTCGGGAG 682
DB 330 LysGlyAsnValGluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGlu 349
QY 683 GCCTATGAGAATGATGTGCTGCCATGAGC-----TTCACCTGATCCTCGCTGCAAC 736
DB 350 SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn 369
QY 737 CTGGCTGCTGAGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCG--CCT 787
DB 370 LeuAsnAlaLeuGlyLeuPheProThrSerGlyMetProProProThrSerGlyPro 389
QY 788 CCAGCAGCGTTACTGGGCTGCTCCTATATAGCTCCTTTATGAGGCTCCCGAGCAGGAG 847
DB 390 ProSerAlaMetThr-----ProProTyrProGlnPheGluGlnSer--GluThrGlu 406
QY 848 ATGCTGCAGGTGTTTCCCGCCAGCAGTGGCGGCATCATCGCAAGAGGGCGAG 907
DB 407 ThrValHisLeuPheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGln 426
QY 908 CACATCAAAACAGCTCTCCCGGTTTCCAGCGCTCCATCAAGATTGCACCCCGAAAACA 967
DB 427 HisIleLysGlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAla 446
QY 968 CTTGACTCCAAAGTTGCTATGTTATCATCTAGCTGACCGCCGAGGCGCCCAATTCAGGCT 1027
DB 447 ProAspAlaLysValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAla 466
QY 1028 CAGGGAAGAATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTG 1087
DB 467 GlnGlyArgGlyIleTyrGlyLysIleLysGluGlnAsnPheValSerProLysGluVal 486
QY 1088 AAGCTGGAGACCCATACGTGTGCCAGCATGAGCAGTGGCGCGGTCATTGCGCAAGGT 1147
DB 487 LysLeuGluAlaHisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGly 506
QY 1148 GAAAAACCGTCAACGAGTGTGCAGAAATTGCGGCGAGCTGAGGTGTTAGTACCAAGAGAC 1207
DB 507 GlyLysThrValAsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgasp 526
QY 1208 CAGACCCCTGTATGAGACGACCCAGCTCATCGTGAATAATCATCGGACATTTCTATGCCAGT 1267
DB 527 GlnThrProAspGluAsnAspGlnValValLysIleThrGlyHisPheTyrAlaCys 546
QY 1268 CAGATGGCTCAACGGAAGATCCGAGACATCTCGGCCAGGTTAAGCAG---CAGCATCAG 1324
DB 547 GlnValAlaGlnArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGln 566
QY 1325 AAGGGA--CAGAGTAACCCAGGCCCGCCAGCAGGAGGAAG 1360
DB 567 LysAlaLeuGlnSerGlyProGlnSerArgArgLys 579

```

```

RESULT 9
US-10-007-700-446
; Sequence 446, Application US/10007700
; Publication No. US20030064947A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Li, Samuel X.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; APPLICANT: Retter, Marc W.
; APPLICANT: Durham, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darriek
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Peckman, David W.
; APPLICANT: Cai, Feng

```



APPLICANT: Fov, Teresa M.  
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER

FILE REFERENCE: 210121.455C17

CURRENT APPLICATION NUMBER: US/10/007,700

CURRENT FILING DATE: 2001-11-30

NUMBER OF SEQ ID NOS: 469

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 446

LENGTH: 579

TYPE: PRT

ORGANISM: Homo sapiens

US-10-007-700-446

Alignment Scores:

| Pred. No.:             | 1.26e-121 | Length:       | 579 |
|------------------------|-----------|---------------|-----|
| Score:                 | 1639.00   | Matches:      | 336 |
| Percent Similarity:    | 83.89%    | Conservative: | 44  |
| Best Local Similarity: | 74.17%    | Mismatches:   | 57  |
| Query Match:           | 52.70%    | Indels:       | 16  |
| DB:                    | 12        | Gaps:         | 9   |

US-09-270-437d-5 (1-1708) x US-10-007-700-446 (1-579)

|    |     |  |     |
|----|-----|--|-----|
| QY | 32  | CGGGAGCATCATGAAGTGAATGGCCACACAGTTGGAGAACCATGCCCTGAAGTCTCC    | 91  |
| DB | 133 | ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla | 152 |
| QY | 92  | TACATCCCGATGACGATAGCA---CAGGACCTGAGATGGCCCGCAGGG-----        | 142 |
| DB | 153 | TyrIleProAspGluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArg    | 172 |
| QY | 143 | GGCTTTGGCTCTCGGGGTCAACCCCGCCAGGGCTCACCTGTGGCAGCGGGGCCACCC    | 202 |
| DB | 173 | GlyLeuGlyGlnArgGlySerSerArgGlnGlySerPro-----GlySerValSer     | 189 |
| QY | 203 | AAGCAGCGAAGTGACATCCCTTCGGCTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCT    | 262 |
| DB | 190 | LysGlnLysProCysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAla | 209 |
| QY | 263 | ATTATTGGCAAGGAGGGGGCCACCATCGCAACATCAAAACAGACCCAGTCCCAAGATA   | 322 |
| DB | 210 | IleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIle | 229 |
| QY | 323 | GACGTGCATAGGAAGGAGAGCGAGTGCAGCTGAGTGAAGGATTCATGATGATGATGATG  | 382 |
| DB | 230 | AspValHisArgLysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThr | 249 |
| QY | 383 | CCTGAGGCTCTCTCGCTGTTAGATGATCTTGGGATTCATGATGATGATGATGATGATG   | 442 |
| DB | 250 | ProGluGlyThrSerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGln | 269 |
| QY | 443 | GACACCAAAACGGGTGACGAGTCTCCCTGAAGATCTTGGCCCAATAAATCTTTGTAGGG  | 502 |
| DB | 270 | AspIleLysPheThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGly | 289 |
| QY | 503 | CGTCTCATGCAAGGAGGAGCGAAGCTGAGAGAGGTAGAGCAAGATACCGAGACAAAA    | 562 |
| DB | 290 | ArgLeuIleGlyLysGluGlyArgAsnLeuLysLysIleGluGlnAspThrAspThrLys | 309 |
| QY | 563 | ATCACCATCTCTCGTTCACAGACTTACCCCTTACCAACCTGAGAGGACCATCAGTGTG   | 622 |
| DB | 310 | IleThrIleSerProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrVal | 329 |
| QY | 623 | AAGGGGCCATCGAAGATTGTTCAGGGCCGAGCAGGAAATATGAAGAAATTCGGGAG     | 682 |
| DB | 330 | LysGlyAsnValGluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGlu    | 349 |
| QY | 683 | GCCTATGAGATGATGTGGCTGCCATGAGC-----TCTCACCTGATCCCTGGCTCAAC    | 736 |
| DB | 350 | SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn | 369 |
| QY | 737 | CTGGTGTGTAGTCTTTTCCCGAGCTTCATCCAGCGAGTCCCGCGG-----CCT        | 787 |

|    |      |  |      |
|----|------|--|------|
| DB | 370  | LeuAsnAlaLeuGlyLeuPheProThrSerGlyMetProProThrSerGlyPro       | 389  |
| QY | 788  | CCAGCAGCGTTACTGGGGCTCTCCCTATAGCTCTCTTATACAGCTCCCGAGGAG       | 847  |
| DB | 390  | ProSerAlaMetThr-----ProProTyrProGlnPheGluGlnSer---GluThrGlu  | 406  |
| QY | 848  | ATGGTCAGGTGTTTATCCCGCCCGAGGCGAGTGGCGCATCATCGGCAAGAGGGCAG     | 907  |
| DB | 407  | ThrValHisLeuPheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGln    | 426  |
| QY | 908  | CACATCAAAACAGCTCTCCCGTTTCCAGCGCTCCATCAAGATTCCACCCCAACA       | 967  |
| DB | 427  | HisIleLysGlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAla | 446  |
| QY | 968  | CCTGACTCCAAAGTTCGTATGTTATCATCTGACCGCCAGCGCCCAATTCAGGCT       | 1027 |
| DB | 447  | ProAspAlaLysValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAla | 466  |
| QY | 1028 | CAGGGAAGAAATCTATGGCAAACTCAAGGAGGAGAACTCTTTGGTCCCAAGGAGGAG    | 1087 |
| DB | 467  | GlnGlyArgIleTyrGlyLysIleLysGluAsnPheValSerProLysGluVal       | 486  |
| QY | 1088 | AGCTGAGACCCACATAGTGTGTCAGCATCAGCAGCTGGCGGCTCTTTGGCAAGGT      | 1147 |
| DB | 487  | LysLeuGluAlaHisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGly | 506  |
| QY | 1148 | GAAAAACCGTCAACAGTTGAGATTTACGCGAGCTGAGTGTGTAGTACCAAGAGAC      | 1207 |
| DB | 507  | GlyLysThrValAsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAsp    | 526  |
| QY | 1208 | CAGACCCCTGATGAGAACGACGAGTCTATCGTGAATAATCATCGACATTTCTATGCA    | 1267 |
| DB | 527  | GlnThrProAspGluAsnAspGlnValValLysIleThrGlyHisPheTyrAlaCys    | 546  |
| QY | 1268 | CAGATGCTCAACGAGATCCGAGACATCTCGCCCGCCAGTTAAGCAG---CAGCATCAG   | 1324 |
| DB | 547  | GlnValAlaGlnArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGln    | 566  |
| QY | 1325 | AAGGGA---CAGAGTAACCGCCAGCGCCAGCGAGGAG 1360                   |      |
| DB | 567  | LysAlaLeuGlnSerGlyProProGlnSerArgArgLys 579                  |      |

RESULT 10

US-10-007-700-449

Sequence 449, Application US/10007700

Publication No. US20030064947A1

GENERAL INFORMATION:

APPLICANT: Wang, Tongtong

APPLICANT: Skeiky, Yasir A.W.

APPLICANT: Li, Samuel X.

APPLICANT: Kalos, Michael D.

APPLICANT: Henderson, Robert A.

APPLICANT: McNeill, Patricia D.

APPLICANT: Fanger, Neil

APPLICANT: Retter, Marc W.

APPLICANT: Durham, Margarita

APPLICANT: Fanger, Gary R.

APPLICANT: Vedvick, Thomas S.

APPLICANT: Carter, Darrick

APPLICANT: Watanabe, Yoshihiro

APPLICANT: Beckman, David W.

APPLICANT: Cai, Feng

APPLICANT: Foy, Teresa M.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER

FILE REFERENCE: 210121.455C17

CURRENT APPLICATION NUMBER: US/10/007,700

CURRENT FILING DATE: 2001-11-30

NUMBER OF SEQ ID NOS: 469

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 449

```

; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-007-700-449

Alignment Scores:
Pred. No.: 1,26e-121 Length: 579
Score: 1639.00 Matches: 336
Percent Similarity: 83.89% Conservative: 44
Best Local Similarity: 74.17% Mismatches: 57
Query Match: 52.70% Indels: 16
DB: 12 Gaps: 9

US-09-270-437D-5 (1-1708) x US-10-007-700-449 (1-579)

QY 32 CGGGAGGACCATCATGAAGCTGAATGGCCACAGATTGGAGACCATGCCCTGAAGTCTCC 91
Db 133 ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla 152
QY 92 TACATCCCGATGACGACATAGCA---CAGGGACCTGAGAAATGGCGCCGAGGG----- 142
Db 153 TyrileProAspGluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArg 172
QY 143 GCGTTTGCTCTCGGGTCACGCCCGCCAGGCTCAGCTGTGGCAGCGGGGCCCGACGCC 202
Db 173 GlyLeuGlyGlnArgGlySerArgGlnGlySerPro-----GlySerValSer 189
QY 203 AAGCAGCAGCAAGTGGACATCCCTTCGGCTCCTGGTGGCCACCCAGTATGTGGGTGCC 262
Db 190 LysGlnLysProCysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAla 209
QY 263 ATTATGCAAGGAGGGGGCCACCATCCGGAACATCAAAACAGACCCAGTCCAGATA 322
Db 210 IleileGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIle 229
QY 323 GACGTGCATAGAGGAGGAGGAGGAGTGCAGCTCAAAAGGCATCAGTGTGCATCCACC 382
Db 230 AspValHisArgLysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThr 249
QY 383 CTGAGGCTCTCTCGCTGTGTAAGATGATCTTGAGATTAATGCAATTAAGAGGCTAAG 442
Db 250 ProGluGlyThrSerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGln 269
QY 443 GACACCAACCGCTGACGAGGTTCCCTGAGATCCTGGCCATATATTAATCTTTAGGG 502
Db 270 AspIleLysPheThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGly 289
QY 503 CGTCTCATTTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 562
Db 290 ArgLeuIleGlyLysGluGlyArgAsnLeuLysLysIleGluGlnAspThrAspThrLys 309
QY 563 ATCACCATCTCTCTGTTGCAAGACTTACCTTTTACACCTGAGGAGGACCATCTGTG 622
Db 310 IleThrIleSerProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrVal 329
QY 623 AAGGGGGCCATCGAGAAATTTGTGACGGCCGAGGAGGAGGAGGAGGAGGAGGAGGAG 682
Db 330 LysGlyAsnValGluThrCysAlaLysAlaGluGluGluIleMetLysLysIleArgGlu 349
QY 683 GCCTATGAGATGATGTGGTGCATGAGC-----TCTCAGCTGATCCCTGCGCTGAAC 736
Db 350 SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn 369
QY 737 CTGCTGCTGTAGTCTTTTCCAGTTCATCCAGCGCATGCCCGCG-----CCT 787
Db 370 LeuAsnAlaLeuGlyLeuPheProProThrSerGlyMetProProProThrSerGlyPro 389
QY 788 CCCAGCAGCGTTACTGGGCTGTCCTATAGTCTCTTTATGAGGCTCCCGAGCAGGAG 847
Db 390 ProSerAlaMetThr-----ProProTyrProGlnPheGluGlnSer---GluThrGlu 406
QY 848 ATGCTGAGGTGTTTATCCCGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 907

```

```

Db 407 ThrValHisLeuPheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGln 426
QY 908 CACATCAAAACAGCTCTCCCGGTTTGGCCAGCGCTCCATCAAGATTGACACCCCAACA 967
Db 427 HisIleLysGlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAla 446
QY 968 CCTGACTCCAAAGTTCGTATGTTATCATCTACCTGACCGCCAGAGGCGCCAAATTCAGGCT 1027
Db 447 ProAspAlaLysValArgMetValIleIleThrGlyProGluAlaGlnPheLysAla 466
QY 1028 CAGGGAGGAATCTATGGCAAACTCAAGGAGGAGAACTTCTTTTGTGCCAAGGAGGAAGT 1087
Db 467 GlnGlyArgIleTyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluVal 486
QY 1088 AAGCTGGAGACCCATACGTGTGCCAGCATCAGCAGCATCAGCGTGGCGGTTCATTGGCAAGGT 1147
Db 487 LysLeuGluAlaHisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGly 506
QY 1148 GGAATAACCGTGAACAGAGTTGCAGAAATTTGACGGCAGCTGAGGTGTAGTACCAAGAGAC 1207
Db 507 GlyLysThrValAsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAsp 526
QY 1208 CAGACCCCTGATGAGAACGACGAGTGCATCTGTGAAATCATCGGACATTTCTATGCCAGT 1267
Db 527 GlnThrProAspGluAsnAspGlnValValLysIleThrGlyHisPheTyrAlaCys 546
QY 1268 CAGATGGCTCAACGGAAGATCCGACACATCTCTGGCCAGGTTAAGCAG---CAGCATCAG 1324
Db 547 GlnValAlaGlnArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGln 566
QY 1325 AAGGGA---CAGAGTAACAGGCCCGCCAGCAGCGAGGAAG 1360
Db 567 LysAlaLeuGlnSerGlyProGlnSerArgArgLys 579

RESULT 11
US-10-117-982-348
; Sequence 348, Application US/10117982
; Publication No. US20030138438A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Mericle, Barbara
; APPLICANT: Spies, Gregory A.
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C18
; CURRENT APPLICATION NUMBER: US/10/117,982
; CURRENT FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 484
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 348
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-117-982-348

Alignment Scores:
Pred. No.: 1,26e-121 Length: 579
Score: 1639.00 Matches: 336
Percent Similarity: 83.89% Conservative: 44
Best Local Similarity: 74.17% Mismatches: 57
Query Match: 52.70% Indels: 16
DB: 14 Gaps: 9

US-09-270-437D-5 (1-1708) x US-10-117-982-348 (1-579)

```

QY 32 CGGGAGCCATCATGAGCTGAATGGCCACAGTTGGAGAACCATGCGCTGAAGGTCTCC 91  
Db 133 ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla 152  
QY 92 TACATCCCGATGACAGATAGCA---CAGGAGCTGAGATGGCCGCCAGGG----- 142  
Db 153 TyrIleProAspGluThrAlaAlaGlnGlnAsnProLeuGlnProArgGlyArg 172  
QY 143 GCGTTTGGCTCTCGGGTACGCCCGCCAGGCTCACCTGGGAGGGGGGCCCGCCAGCC 202  
Db 173 GlyLeuGlnArgGlySerArgGlnGlySerPro-----GlySerValSer 189  
QY 203 AAGCAGCAGCAAGTGGCATCCCTCGGCTCTCGGTGGCCACCCAGTATGGGTGCC 262  
Db 190 LysGlnLysProCysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAla 209  
QY 263 ATTATTGGCAGGAGGGGGCCACCATCCGCAACATCAAAACAGACCCAGTCCAGATA 322  
Db 210 IleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIle 229  
QY 323 GACGTGCATAGAGGAGAGCGAGGTGCAGCTGAAAAAGCCATCAGTGTGCATCCACC 382  
Db 230 AspValHisArgLysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThr 249  
QY 383 CCGTGGGCTGCTCTCGCTTGTGAAGATGATCTTGGAGATATGCTAATAAGAGCTAAG 442  
Db 250 ProGluGlyThrSerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGln 269  
QY 443 GACACCAAACGGCTGACAGGTCCCTGAGATCTTGGCCCATATAACTTTGTAGGG 502  
Db 270 AspileLysPheThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGly 289  
QY 503 CGTCTCATGGCAAGGAGGACGAACTGAGAGGTAGAGCAAGATACCGAGACAAA 562  
Db 290 ArgLeuIleGlyLysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLys 309  
QY 563 ATCACCATCTCTGTTGCAAGCCTTACACCTTTACACCTGAGAGACCATCACTGTG 622  
Db 310 IleThrIleSerProLeuGlnGluLeuThrLeuTyrrAsnProGluArgThrIleThrVal 329  
QY 623 AAGGGGGCCATCGAGAATTTGTGAGGGCCGAGCAGGAGAAATAATGAAGAAAGTTCGGGAG 682  
Db 330 LysGlyAsnValGluThrCysAlaLysAlaGluGluGluIleMetLysLysIleArgGlu 349  
QY 683 GCCTATGAGATGATGTGGTCCCATGAGC-----TCTACCTGATCCCTGGCTGAAC 736  
Db 350 SerTyrrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn 369  
QY 737 CTGGCTGTGTAGGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCG-----CCT 787  
Db 370 LeuAsnAlaLeuGlyLeuPheProThrSerGlyMetProProProThrSerGlyPro 389  
QY 788 CCCAGCAGGTACTGGGGCTGCTCCCTATAGCTCTCTTATGACAGGTCCCGAGCAGGAG 847  
Db 390 ProSerAlaMetThr-----ProProTyrrProGlnPheGluGlnSer---GluThrGlu 406  
QY 848 ATGGTGCAGGTGTTATCCCGCCAGCAGTGGGCGCATCCGCAAGAGGGGCGAG 907  
Db 407 ThrValHisLeuPheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGln 426  
QY 908 CACATCAACAGCTCTCCGGTTTCCAGCGCTCCATCAAGATTGCAACCCCGCAACA 967  
Db 427 HisIleLysGlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAla 446  
QY 968 CTGACTCCAAAGTGTGATGTTATCATCTAGCAGCGCCGAGGCGCCCAATTCAGGCT 1027  
Db 447 ProAspAlaLysValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAla 466  
QY 1028 CAGGGAAGAACTATGAGCAACTCAAGGAGGAGAACTCTTTGGTCCCAAGGAGGAAGTG 1087  
Db 467 GlnGlyArgIleTyrrGlyLysIleLysGluGluAsnPheValSerProLysGluGluVal 486  
QY 1088 AAGCTGGAGACCCCATACGTGTGCCATCAGCAGGTGGCGGGTCAATTGGCAAGGT 1147

Db 487 LysLeuGluAlaHisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGly 506  
QY 1148 GGAATAACCGTGAACAGTTTCAGATTTTCACGGCAGCTGAGGTGGTAGTACCAAGAC 1207  
Db 507 GlyLysThrValAsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgasp 526  
QY 1208 CAGACCCCTGATGAGAACCGACCGAGTCTCGTCAAAATCATCCGACATTTCTATGCCAGT 1267  
Db 527 GlnThrProAspGluAsnAspGlnValValLysIleThrGlyHisPheTyrrAlaCys 546  
QY 1268 CAGATGGCTCAACGGAGATCCGAGACATCTCTGGCCAGGTAAAGCAG---CAGCATCAG 1324  
Db 547 GlnValAlaGlnArgLysIleGlnIleLeuThrGlnValLysGlnHisGlnGlnGln 566  
QY 1325 AAGGGA---CAGAGTAACCGACGCCAGGCGACGAGGAAG 1360  
Db 567 LysAlaLeuGlnSerGlyProGlnSerArgArgLys 579

## RESULT 12

US-10-117-982-446  
; Sequence 446, Application US/10117982  
; Publication No. US20030138438A1  
; GENERAL INFORMATION:  
; APPLICANT: Foy, Teresa M.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Mericle, Barbara  
; APPLICANT: Spies, Gregory A.  
; APPLICANT: Fan, Ligu  
; APPLICANT: Wang, Tongtong  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.455C18  
; CURRENT APPLICATION NUMBER: US/10/117,982  
; CURRENT FILING DATE: 2002-04-05  
; NUMBER OF SEQ ID NOS: 484  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 446  
; LENGTH: 579  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-117-982-446

Alignment Scores:  
Pred. No.: 1-26e-121 Length: 579  
Score: 1639.00 Matches: 336  
Percent Similarity: 83.89% Conservative: 44  
Best Local Similarity: 74.17% Mismatches: 57  
Query Match: 52.70% Indels: 16  
DB: 14 Gaps: 9

US-09-270-437D-5 (1-1708) x US-10-117-982-446 (1-579)

QY 32 CGGGAGCCATCATGAGCTGAATGGCCACAGTTGGAGAACCATGCGCTGAAGGTCTCC 91  
Db 133 ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla 152  
QY 92 TACATCCCGATGACAGATAGCA---CAGGAGCTGAGATGGCCGCCAGGG----- 142  
Db 153 TyrIleProAspGluThrAlaAlaGlnGlnAsnProLeuGlnProArgGlyArg 172  
QY 143 GCGTTTGGCTCTCGGGTACGCCCGCCAGGCTCACCTGGGAGGGGGGCCCGCCAGCC 202  
Db 173 GlyLeuGlnArgGlySerArgGlnGlySerPro-----GlySerValSer 189  
QY 203 AAGCAGCAGCAAGTGGCATCCCTCGGCTCTCGGTGGCCACCCAGTATGGGTGCC 262  
Db 190 LysGlnLysProCysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAla 209

QY 263 ATTATTGCAAGAGGGGCCACCATCGCAACATCAAAAACAGACCCACATCCCAAGATA 322  
 Db 210 llelleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIle 229  
 QY 323 GACGTGCATAGGAAGAGCAACGACGTGCAGCTGAAAAAGCCATCAGTGTGCATCTCCACC 382  
 Db 230 AspValHisArgLysGluAsnAlaGlyAlaGluLysSerIleThrIleLeuSerThr 249  
 QY 383 CTGAGGGCTGCTCTCCGCTGTGAAGATGATTTGGAGATTATGCATAAAGAGCGCTAAG 442  
 Db 250 ProGluGlyThrSerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGln 269  
 QY 443 GACACCAAAAACGGCTGACAGGTTCCCTGAGATCTCGGCCCATATATCTTTGTAGGG 502  
 Db 270 AspileLysPheThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGly 289  
 QY 503 GTCTCATTTGCAAGAGGACGGAACCTGAAGAGGTAGACGAAGATACCGACACAAA 562  
 Db 290 ArgLeuIleGlyLysGluGlyArgAsnLeuLysLysIleGluGlnAspThrAspThrLys 309  
 QY 563 ATCCATCTCTCGTTCAGACCTTACCTTTACACCTGAGACCATCAGACCATCTGTG 622  
 Db 310 lleThrIleSerProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrVal 329  
 QY 623 AAGGGGGCCATCGAATTTTCCAGCTTCATCCAGCGCAGTCCGCGCG 682  
 Db 330 LysGlyAsnValGluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGlu 349  
 QY 683 GCCTATGAGATGATGTGCTGCCATGAGC-----TCTCAGCTGATCCCTGGCCTGAAC 736  
 Db 350 SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn 369  
 QY 737 CTGGTGTGTAGGTTCTTTCCAGCTTCATCCAGCGCAGTCCGCGCG-----CCT 787  
 Db 370 LeuAsnAlaLeuGlyLeuPheProProThrSerGlyMetProProThrSerGlyPro 389  
 QY 788 CCCAGACGGTTACTGGGGCTGCTCCCTATAGCTCTTTATGACGGTCCCGAGCAGGAG 847  
 Db 390 ProSerAlaMetThr-----ProProTyrProGlnPheGluGlnSer---GluThrGlu 406  
 QY 848 ATGTGTCAGGTGTTATCCCGCCAGCAGTGGCGCCCATCATCCGCAAGAGGGGCGAG 907  
 Db 407 ThrValHisLeuPheIleProAlaLeuSerValGlyAlaIlelleGlyLysGlnGlyGln 426  
 QY 908 CACATCAAAACAGCTCTCCCGGTTTCCAGCGCCTCCATCAAGATTGCAACCCCGAACA 967  
 Db 427 HisIleLysGlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAla 446  
 QY 968 CTGACTCCAAAGTTCGTATGGTTATCATCTAGGACCGCCAGAGGCCCAATTCAGGCT 1027  
 Db 447 ProAspAlaLysValA-gmetValIleIleThrGlyProGluAlaGlnPheLysAla 466  
 QY 1028 CAGGCAAGATCTATGGCAAACTCAAGGAGAGAACTTCTTTGTCCTCAAGAGGAGTG 1087  
 Db 467 GlnGlyArgIleTyrGlyLysIleLysGluLysPheValSerProLysGluGluVal 486  
 QY 1088 AAGCTGAGAGCCACATACGTGTGCCAGCATCAGACGTGGCGGGTCAATGCAAGGT 1147  
 Db 487 LysLeuGluAlaHisIleA-gvalProSerPheAlaAlaGlyArgValIleGlyLysGly 506  
 QY 1148 GGA AAAACGGTGAACGAGTTGCAATTTGACGCGCAGCTCAGGTGTAGTACCAAGAGAC 1207  
 Db 507 GlyLysThrValAsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAsp 526  
 QY 1208 CAGACCCCTGATGAGAACGACAGGTGCATCGTAAAAATCATCCGACATTTCTATGCCAGT 1267  
 Db 527 GlnThrProAspGluAsnAspGlnValValLysIleThrGlyHisPheTyrAlaCys 546  
 QY 1268 CAGATGCTCAACGGAAGATCCGAGCATCTCGGCCAGGTTAAGCAG---CAGCATCAG 1324  
 Db 547 GlnValAlaGlnArgLysIleGlnIleLeuThrGlnValLysGlnHisGlnGlnGln 566

QY 1325 AAGGGA---CAGAGTAAACAGCCCGCCAGGACCGAGAGAG 1360  
 Db 567 LysAlaLeuGlnSerGlyProGlnSerArgArgLys 579  
 RESULT 13  
 US-10-117-982-449  
 ; Sequence 449, Application US/10117982  
 ; Publication No. US20030138438A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Foy, Teresa M.  
 ; APPLICANT: Fanger, Gary R.  
 ; APPLICANT: Vedvick, Thomas S.  
 ; APPLICANT: Carter, Darrick  
 ; APPLICANT: Watanabe, Yoshihiro  
 ; APPLICANT: Henderson, Robert A.  
 ; APPLICANT: Kalos, Michael D.  
 ; APPLICANT: Mericle, Barbara  
 ; APPLICANT: Spies, Gregory A.  
 ; APPLICANT: Fan, Licun  
 ; APPLICANT: Wang, Tonglong  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
 ; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
 ; FILE REFERENCE: 210121.455C18  
 ; CURRENT APPLICATION NUMBER: US/10/117,982  
 ; CURRENT FILING DATE: 2002-04-05  
 ; NUMBER OF SEQ ID NOS: 484  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 449  
 ; LENGTH: 579  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-117-982-449

Alignment Scores:  
 Pred. No.: 1,26e-121 Length: 579  
 Score: 1639.00 Matches: 336  
 Percent Similarity: 83.89% Conservative: 44  
 Best Local Similarity: 74.17% Mismatches: 57  
 Query Match: 52.70% Indels: 16  
 DB: 14 Gaps: 9  
 US-09-270-437d-5 (1-1708) x US-10-117-982-449 (1-579)  
 QY 32 CGGGAGCCATCATGAAGCTGAATGGCCACCATGGAGAACCATGCCCTGAAGGTCTCC 91  
 Db 133 ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla 152  
 QY 92 TACATCCCGATGACAGCATAGCA---CAGGACCTGAGATGGCGCCGAGGG----- 142  
 Db 153 TyrIleProAspGluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArg 172  
 QY 143 GCGTTTGGCTCTCGGGGTGACGCGCCCGCCAGGGCTCACCTGTGGCAGCGGGGGCCAGCC 202  
 Db 173 GlyLeuGlyGlnArgGlySerSerArgGlnGlySerPro-----GlySerValSer 189  
 QY 203 AAGCAGACGAAAGTGGAGCATCCCTCTCGCTCTCTGCTGCCACCCAGTATGTGGGTGCC 262  
 Db 190 LysGlnLysProCysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAla 209  
 QY 263 ATTATTTGCAAGAGGGGGCCACCATCCGCAACATCAAAAACAGACCCAGTCCCAAGATA 322  
 Db 210 llelleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIle 229  
 QY 323 GACGTGCATAGGAAGAGGACGACGAGTGCAGCTGAGGAGGAGGAGGAGGAGGAGGAGG 382  
 Db 230 AspValHisArgLysGluAsnAlaGlyAlaGluLysSerIleThrIleLeuSerThr 249  
 QY 383 CCTGAGGGCTGCTCTCGCTTGTGAAGATGATCTTGGAGATTATGCATAAAGAGGCTAAG 442  
 Db 250 ProGluGlyThrSerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGln 269  
 QY 443 CACACCAAAACGGTGCAGAGGTTCCCTGAGATCTCGGCCAGGATAAATCTTTGTAGGG 502

Db 270 AspileLysPheThrGluGluileProLeuLysileLeuAlaHisAsnAsnPheValGly 289  
Qy 503 CGTCTATTGGCAAGGACCGAACCTCAAGAGGTAGACAGATACCGAGACAAA 562  
Db 290 ArgLeuileGlyLysGluGlyArgAsnLeuLysileGluGlnAspThrLys 309  
Qy 563 ATACCATCTCCCTGCTTGCAGACCTTACCCCTTTACAACTCAGAGCCATCACTGTG 622  
Db 310 IleThrileSerProLeuGlnGluLeuThrLeuTyAsnProGluArgThrileThrVal 329  
Qy 623 AAGGGGCCATCGAAGATTGTCAGGCGCGAGCAGGAGAAATATAGAAAGTTCGGAG 682  
Db 330 LysGlyAsnValGluThrCysAlaLysAlaGluGluileMetLysLysileArgGlu 349  
Qy 683 GCCTATCAGATGATGGTGCATGAGC-----TCTCACCTGATCCCTGCGCTGAAC 736  
Db 350 SerTyGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisleulleProGlyLeuAsn 369  
Qy 737 CTGGCTCTGTAGTCTTTTCCAGCTTTCATCCAGCGCAGTCCCGCG-----CCT 787  
Db 370 LeuAsnAlaLeuGlyLeuPheProProThrSerGlyMetProProThrSerGlyPro 389  
Qy 788 CCCAGCAGCTTACTGGGGTGTCTCCCTATAGCTCTTTATGAGGCTCCCGAGCAGGAG 847  
Db 390 ProSerAlaMetThr-----ProProTyProGlnPheGluGlnSer---GluThrGlu 406  
Qy 848 ATGGTCAGGTGTTTATCCCGCCAGCGCAGTGGCGCCATCATCGGCAAGAGGGCGAG 907  
Db 407 ThrValHisLeuPheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGln 426  
Qy 908 CACATCAAAACGCTCTCCGGTTGCGAGCGCTCCATCAAGATTGACACCCGCAACA 967  
Db 427 HisileLysGlnLeuSerArgPheAlaGlyAlaSerileLysileAlaProAlaGluAla 446  
Qy 968 CTTGACTCCAAAGTTGTTGTTATCATCTGACGCGCGCAGAGGCCCAATCAAGGCT 1027  
Db 447 ProAspAlaLysValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAla 466  
Qy 1028 CAGGGAAGAACTTATGCAAACTCAAGAGGAGAACTTCTTTGTCCTCAAGAGGAGAGTG 1087  
Db 467 GlnGlyArgIleTyGlyLysileLysGluGluLeuAsnPheValSerProLysGluGluVal 486  
Qy 1088 AAGCTGAGACCCACATACGTGTCAGCATCAGCATCAGCATCGCGGGTCATTGGCAAGGT 1147  
Db 487 LysLeuGluAlaHisileArgValProSerPheAlaAlaGlyArgValIleGlyLysGly 506  
Qy 1148 GGAAAAACGGTGAACGAGTTTCAGAAATTTGACGGCAGCTGAGGTGGTAGTACCAAGAGAC 1207  
Db 507 GlyLysThrValAsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAsp 526  
Qy 1208 CAGACCCCTGATGAGAACGACGAGTGCATCGTGAATATCATCGACATTTCTATGCCAGT 1267  
Db 527 GlnThrProAspGluAsnAspGlnValValValLysileThrGlyHisPheTyAlaCys 546  
Qy 1268 CAGATGGCTCAACGGAAGATCCGAGACATCTGCGCCAGGTAAAGCAG---CAGCATCAG 1324  
Db 547 GlnValAlaGlnArgLysileGlnGluileGlnLeuThrGlnValLysGlnHisGlnGln 566  
Qy 1325 AAGGGA---CAGAGTAAACCGCCAGCGCAGCAGGAGGAAG 1360  
Db 567 LysAlaLeuGlnSerGlyProProGlnSerArgLys 579

RESULT 14  
US-10-117-982-480  
; Sequence 480, Application US/10117982  
; Publication No. US20030138438A1  
; GENERAL INFORMATION:  
; APPLICANT: Foy, Teresa M.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Vedwick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Henderson, Robert A.

; APPLICANT: Kalos, Michael D.  
; APPLICANT: Mericle, Barbara  
; APPLICANT: Spies, Gregory A.  
; APPLICANT: Fan, Liqun  
; APPLICANT: Wang, Tongcong  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.455C18  
; CURRENT APPLICATION NUMBER: US/10/117,982  
; CURRENT FILING DATE: 2002-04-05  
; NUMBER OF SEQ ID NOS: 484  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 480  
; LENGTH: 579  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-117-982-480

Alignment Scores:  
Pred. No.: 1,268-121 Length: 579  
Score: 1639.00 Matches: 336  
Percent Similarity: 83.89% Conservative: 44  
Best Local Similarity: 74.17% Mismatches: 57  
Query Match: 52.70% Indels: 16  
DB: 14 Gaps: 9

US-09-270-437D-5 (1-1708) x US-10-117-982-480 (1-579)

Qy 32 CGCGGAGCCATCATGAAGCTGAATCGCCACAGTTGGAGAACCATGCTGAGGTCTCC 91  
Db 133 ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnPheThrileLysValAla 152  
Qy 92 TACATCCCGATGAGCAGATAGCA---CAGGGACCTGAGAAATGGCGCGGAGGG----- 142  
Db 153 TyrIleProAspGluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArg 172  
Qy 143 GCCTTTGGCTCTCGGGGTCAGCCCGCCAGCGCTCCTGCTGGCAGCGGGGCCCGAGCC 202  
Db 173 GlyLeuGlyLysArgGlySerSerArgGlnGlySerPro-----GlySerValSer 189  
Qy 203 AAGCAGCAGCAAGTAGTGACATCCCCCTTCGGCTCTCGGTGGTGGCCACCCAGTATGTGGTGCC 262  
Db 190 LysGlnLysProCysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAla 209  
Qy 263 ATTATTGGCAAGAGGGGGCCACCATCCGACATCACAACACAGACAGCCAGTCCCAAGATA 322  
Db 210 IleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysile 229  
Qy 323 GACGTGCATAGGAAGAGAGAACGAGGTGAGTGTGAAAAAGCCATCAGTGTGCATCTCCACC 382  
Db 230 AspValHisArgLysGluAsnAlaGlyAlaAlaGluLysSerIleThrileLeuSerThr 249  
Qy 383 CTGAGGGTCTCTCCCTGTTAAGATGATCTTGGAGATTCATGATTAAGAGGCTAAG 442  
Db 250 ProGluGlyThrSerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGln 269  
Qy 443 GACACCAACCGCTGACGAGGTTCCTGAGAGATCCTGGCCATAATAAATCTTGTAGGG 502  
Db 270 AspIleLysPheThrGluGluileProLeuLysileLeuAlaHisAsnAsnPheValGly 289  
Qy 503 CECTCTATTGGCAAGGAGCGAACCTGAAGAAGTAGACACAGATACCGAGACAAA 562  
Db 290 ArgLeuIleGlyLysGluGlyArgAsnLeuLysLysileGluGlnAspThrAspThrLys 309  
Qy 563 ATCCCATCTCTCTGTTGCAAGACCTTACCTTACACCTGAGAGGACCATCACTG 622  
Db 310 IleThrileSerProLeuGlnGluLeuThrLeuTyAsnProGluArgThrileThrVal 329  
Qy 623 AAGGGGCCATCGAAGATTGTCAGGCGCGAGGAGAAATATAGAAAGTTCGGAG 682  
Db 330 LysGlyAsnValGluThrCysAlaLysAlaGluGluileMetLysLysileArgGlu 349  
Qy 683 GCCTATGAGATGATGGTGCATGAGC-----TCTCACCTGATCCCTGCGCTGAAC 736

```

Db      350 SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn 369
QY      737 CTGGCTGCTGAGGCTTTTCCAGCTTCATCCAGCCGAGTCCGCCG-----CCT 787
Db      370 LeuAsnAlaLeuGlyLeuPheProThrSerGlyMetProProThrSerGlyPro 389
QY      788 CCCAGCAGGTTACTGGGCTGCTCCCTATAGCTCTTTATGCGAGGCTCCCGCAGGAG 847
Db      390 ProSerAlaMetThr-----ProProTyrProGlnPheGluGlnSer---GluThrGlu 406
QY      848 ATGTGTGAGGTGTTTATCCCGCCGAGCAGTGGCGCCCATCATCGGCAAGAGGGGCGAG 907
Db      407 ThrValHisLeuPheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGln 426
QY      908 CACATCAACAGCTCTCCCGGTTTCCAGCGCCCTCCATCAAGATTGACACCCCGAACA 967
Db      427 HisIleLysGlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAla 446
QY      968 CTTGACTCCAAAGTTTGTATGTTATCATCTGATCGACCGCCAGAGGCCCAATTCAAGGCT 1027
Db      447 ProAspAlaLysValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAla 466
QY      1028 CAGGGAAGATCTATGGCAACTCAAGAGAGAGAACTCTTTTGGTCCCAAGGAGGAAGTG 1087
Db      467 GlnGlyArgIleTyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluVal 486
QY      1088 AAGCTGGAGACCCACATACGTGTCCAGCATCAGCAGCTGGCGGGTCATTGGCAAGGT 1147
Db      487 LysLeuGluAlaHisIleArgValProSerPheAlaGlyArgValIleGlyLysGly 506
QY      1148 GGAAAAAGCGTGAACGAGTTGCAGAAATTGACGCGACCTGAGTGTGAGTACCAAGAGAC 1207
Db      507 GlyLysThrValAsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAsp 526
QY      1208 CAGACCCCTGATGAGACGACCGAGTCTCATCGTGAATCATCGACATCTTTCTATGCCAGT 1267
Db      527 GlnThrProAspGluAsnAspGlnValValLysIleThrGlyHisPheTyrAlaCys 546
QY      1268 CAGATGCTCAACGGAAGATCCGAGACATCTTGCCCGCAGGTTAAGCAG---CAGCATCAG 1324
Db      547 GlnValAlaGlnArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGln 566
QY      1325 AAGSGA---CAGATTAACCGCCCGCAGGCGAGGAG 1360
Db      567 LysAlaLeuGlnSerGlyProProGlnSerArgArgLys 579

```

## RESULT 15

```

US-10-313-986-348
; Sequence 348, Application US/10313986
; Publication No. US20030236209A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: McNabb, Andria
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C19
; CURRENT APPLICATION NUMBER: US/10/313,986
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 348
; LENGTH: 579
; TYPE: PRN
; ORGANISM: Homo sapiens
US-10-313-986-348

```

Alignment Scores:

```

Pred. No.: 1,26e-121 Length: 579
Score: 1639.00 Matches: 336

```

```

Percent Similarity: 83.89% Conservative: 44
Best Local Similarity: 74.17% Mismatches: 57
Query Match: 52.70% Indels: 16
DB: 15 Gaps: 9
US-09-270-437D-5 (1-1708) x US-10-313-986-348 (1-579)
QY      32 CGGGAGGACCATCAAGCTGAATGCCACCATGTTGGAGAACCATGCCCTCAAGGCTCTCC 91
Db      133 ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla 152
QY      92 TACATCCCCGATGACGACAGATAGCA---CAGGAGACCTGAGAATGGCGCCGAGGG----- 142
Db      153 TyrIleProAspGluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArg 172
QY      143 GGCCTTTGGCTCTCGGGGTACGCCCGCCAGGGCTCACCTGTGGCAGCGGGGGCCCGACCC 202
Db      173 GlyLeuGlyGlnArgGlySerSerArgGlnGlySerPro-----GlySerValSer 189
QY      203 AAGCAGCAGCAAGTGCACATCCCTTCGGCTCCTGTGGTCCCGCCACCCAGTATGTGGTGCC 262
Db      190 LysGlnLysProCysAspLeuProLeuArgLeuValProThrGlnPheValGlyAla 209
QY      263 ATTATTGCAAGGAGGGGGCCACCATCCGCAACATCAAAACACAGACCCAGCTCCAAAGATA 322
Db      210 IleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIle 229
QY      323 GACGTGCATAGGAGGAGACGCGAGTGCAGTCAAGAAAGCCATCAGTGTGCATCCACCC 382
Db      230 AspValHisArgLysGluAsnAlaGlyAlaGluLysSerIleThrIleLeuSerThr 249
QY      383 CTTGAGGGCTGCTCTCCGCTTGTAAAGATGATCTTTGGAGATTATGCAATAAGAGGCTAAG 442
Db      250 ProGluGlyThrSerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGln 269
QY      443 GACACCAAGCGGTGACGAGGTTCCTCTGAAGATCTCGGCCCATATACTTTGTAGGG 502
Db      270 AspileLysPheThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGly 289
QY      503 CTTCTCATTTGCAAGAGGACGAACTGAAAGAGGTAGAGCAAGATACCGACACAAA 562
Db      290 ArgLeuIleGlyLysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLys 309
QY      563 ATCACCATTCTCTGTCGAAGACCTTACCTTTTACAACTGAGAGAGACCATCACTGTG 622
Db      310 IleThrIleSerProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrVal 329
QY      623 AAGGGGCCATCGAGAATTGTTGAGGGCCGAGCAGGAAATTAATGAAGAAGTTCCGGAG 682
Db      330 LysGlyAsnValGluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGlu 349
QY      683 GCCTATGAGAATGATGTGGTGCATCAGC-----TCTCACCTGATCCCTGGCGCTGAAC 736
Db      350 SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn 369
QY      737 CTGGCTGCTGAGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCCG-----CCT 787
Db      370 LeuAsnAlaLeuGlyLeuPheProThrSerGlyMetProProThrSerGlyPro 389
QY      788 CCCAGCAGGTTTACTGGGGCTGCTCCCTATAGCTCTTTATGCGAGGCTCCCGCAGGAG 847
Db      390 ProSerAlaMetThr-----ProProTyrProGlnPheGluGlnSer---GluThrGlu 406
QY      848 ATGTGTGAGGTGTTTATCCCGCCGAGCAGTGGCGCCCATCATCGGCAAGAGGGGCGAG 907
Db      407 ThrValHisLeuPheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGln 426
QY      908 CACATCAACAGCTCTCCCGGTTTCCAGCGCCCTCCATCAAGATTGACACCCCGAACA 967
Db      427 HisIleLysGlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAla 446
QY      968 CTTGACTCCAAAGTTTGTATGTTATCATCTGATCGACCGCCAGAGGCCCAATTCAAGGCT 1027

```

Db 447 ProAspAlaLysValArgMetValIleIleThrGlyProGluAlaGlnPheLysAla 466  
Qy 1028 CAGGGAAGAAATCTATGCAAACTCAAGGAGGAGAACTTCTTGTCCCAAGGAGGAAGTG 1087  
Db 467 GlnGlyArgIleTyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluVal 486  
Qy 1088 AAGCTGAGACCCACATACGTGTGCCAGCATCAGACGTGGCCGGTCAATTGGCAAAGGT 1147  
Db 487 LysLeuGluAlaHisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGly 506  
Qy 1148 GGAATAACGGTGAACGAGTTTCAGAAATTTGACCGCAGCTGAGGTGGTAGTACCAAGAGAC 1207  
Db 507 GlyLysThrValAsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgASP 526  
Qy 1208 CAGACCCCTGATGAGAAACGACCAAGTCAATCGTGAAAAATCATCGGACATTTCTATGCCAGT 1267  
Db 527 GlnThrProAspGluAsnAspGlnValValLysIleThrGlyHisPheTyrAlaCys 546  
Qy 1268 CAGATGGCTCAACGGAAGATCCGAGACATCTCTGCCCCAGGTTAAGCAG---CAGCATCAG 1324  
Db 547 GlnValAlaGlnArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGln 566  
Qy 1325 AAGGGA---CAGAGTAACCAAGGCCAGGCACGAGGAAG 1360  
Db 567 LysAlaLeuGlnSerGlyProGlnSerArgArgLys 579

Search completed: July 13, 2004, 12:27:58  
Job time : 110 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model  
Run on: July 13, 2004, 12:28:03 ; Search time 22 seconds  
(without alignments)  
8016.105 Million cell updates/sec

Title: US-09-270-437D-5  
Perfect score: 580  
Sequence: 1 agggagctgcgcacccgc.....attctctcaggttttaaaa 1708

Scoring table: OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 389414 seqs, 51625971 residues  
Word size: 1

Total number of hits satisfying chosen parameters: 663654  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:  
-MODE=frame+ n2p.model -DEV=xlp  
-Q=/cgn2\_1/uspro.spool p/US09270437/runat 13072004.121956.10024/app query.fasta\_1.1863  
-DB=Issued Patents AA -OPMT=fastan -SUFFIX=rai -MINMATCH=0 1 -LOOPCI=0  
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR.SCORE=quality -THR.MIN=1 -ALIGN=15 -MODE=LOCAL  
-OUTPMT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09270437 @cgn 1\_1\_33 @runat 13072004.121956.10024 -NCFU=6 -ICPU=3  
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PTUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

| SUMMARIES  |       |             |        |                    |                   |
|------------|-------|-------------|--------|--------------------|-------------------|
| Result No. | Score | Query Match | Length | ID                 | Description       |
| 1          | 166   | 29.6        | 577    | US-09-261-855-2    | Sequence 2, Appli |
| 2          | 49    | 8.8         | 49     | US-09-261-855-18   | Sequence 18, Appl |
| 3          | 48    | 8.6         | 48     | US-09-261-855-20   | Sequence 20, Appl |
| 4          | 47    | 8.4         | 47     | US-09-261-855-17   | Sequence 17, Appl |
| 5          | 45    | 8.0         | 47     | US-09-261-855-19   | Sequence 19, Appl |
| 6          | 36    | 6.4         | 579    | US-09-643-597-176  | Sequence 176, App |
| 7          | 36    | 6.4         | 579    | US-09-643-597-348  | Sequence 348, App |
| 8          | 36    | 6.4         | 579    | US-09-480-884A-176 | Sequence 176, App |
| 9          | 36    | 6.4         | 579    | US-09-542-615A-176 | Sequence 176, App |
| 10         | 36    | 6.4         | 579    | US-09-542-615A-348 | Sequence 348, App |
| 11         | 36    | 6.4         | 579    | US-09-606-421B-176 | Sequence 176, App |
| 12         | 36    | 6.4         | 579    | US-09-606-421B-348 | Sequence 348, App |

|    |    |     |      |   |                      |                    |
|----|----|-----|------|---|----------------------|--------------------|
| 13 | 28 | 5.0 | 47   | 3 | US-09-261-855-21     | Sequence 21, Appl  |
| 14 | 19 | 3.4 | 48   | 3 | US-09-261-855-24     | Sequence 24, Appl  |
| 15 | 14 | 2.5 | 14   | 3 | US-09-261-855-3      | Sequence 3, Appli  |
| 16 | 14 | 2.5 | 14   | 3 | US-09-261-855-46     | Sequence 46, Appl  |
| 17 | 14 | 2.5 | 49   | 3 | US-09-261-855-22     | Sequence 22, Appl  |
| 18 | 11 | 2.0 | 11   | 3 | US-09-261-855-13     | Sequence 13, Appl  |
| 19 | 11 | 2.0 | 47   | 3 | US-09-261-855-23     | Sequence 23, Appl  |
| 20 | 9  | 1.6 | 529  | 4 | US-09-381-656-1      | Sequence 1, Appli  |
| 21 | 8  | 1.4 | 142  | 4 | US-09-252-991A-23673 | Sequence 23673, A  |
| 22 | 8  | 1.4 | 161  | 4 | US-09-252-991A-24938 | Sequence 24938, A  |
| 23 | 8  | 1.4 | 212  | 4 | US-09-252-991A-32417 | Sequence 32417, A  |
| 24 | 8  | 1.4 | 287  | 4 | US-09-252-991A-29951 | Sequence 29951, A  |
| 25 | 8  | 1.4 | 317  | 4 | US-09-252-991A-30984 | Sequence 30984, A  |
| 26 | 8  | 1.4 | 322  | 4 | US-09-252-991A-25412 | Sequence 25412, A  |
| 27 | 8  | 1.4 | 448  | 4 | US-09-543-681A-6550  | Sequence 6550, Ap  |
| 28 | 8  | 1.4 | 2568 | 4 | US-09-866-108A-3     | Sequence 3, Appli  |
| 29 | 7  | 1.3 | 21   | 6 | 5368712-7            | Patent No. 5368712 |
| 30 | 7  | 1.3 | 23   | 3 | US-09-093-227-2      | Sequence 2, Appli  |
| 31 | 7  | 1.3 | 30   | 3 | US-09-136-251-7      | Sequence 7, Appli  |
| 32 | 7  | 1.3 | 30   | 4 | US-09-634-496-7      | Sequence 7, Appli  |
| 33 | 7  | 1.3 | 30   | 4 | US-09-635-145A-7     | Sequence 7, Appli  |
| 34 | 7  | 1.2 | 34   | 4 | US-09-079-030-57     | Sequence 57, Appl  |
| 35 | 7  | 1.2 | 35   | 4 | US-08-722-015A-224   | Sequence 224, App  |
| 36 | 7  | 1.2 | 50   | 3 | US-09-261-855-25     | Sequence 25, Appl  |
| 37 | 7  | 1.3 | 55   | 3 | US-09-162-934-14     | Sequence 14, Appl  |
| 38 | 7  | 1.3 | 56   | 4 | US-09-187-789-64     | Sequence 64, Appl  |
| 39 | 7  | 1.3 | 56   | 4 | US-09-139-600-59     | Sequence 59, Appl  |
| 40 | 7  | 1.3 | 57   | 1 | US-08-137-614A-19    | Sequence 19, Appl  |
| 41 | 7  | 1.3 | 63   | 4 | US-08-963-851-18     | Sequence 18, Appl  |
| 42 | 7  | 1.2 | 78   | 4 | US-09-540-236-2668   | Sequence 2668, Ap  |
| 43 | 7  | 1.2 | 85   | 4 | US-09-732-210-1119   | Sequence 1119, Ap  |
| 44 | 7  | 1.2 | 90   | 4 | US-09-040-229B-12    | Sequence 12, Appl  |
| 45 | 7  | 1.3 | 97   | 4 | US-09-621-976-5598   | Sequence 5598, Ap  |

ALIGNMENTS

RESULT 1  
US-09-261-855-2  
; Sequence 2, Application US/09261855A  
; Patent No. 6255055  
; GENERAL INFORMATION:  
; APPLICANT: Ross, Jeffrey  
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN  
; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE  
; FILE REFERENCE: 960296.95131  
; CURRENT APPLICATION NUMBER: US/09/261,855A  
; CURRENT FILING DATE: 1999-03-03  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 577  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-261-855-2

|   |          |   |     |
|---|----------|---|-----|
| Alignment Scores:                                   | 6.4e-156 | Length:   | 577 |
| Pred. No.:  | 166.00   | Matches:  | 166 |
| Score:  | 100.00%  | Conservative:   | 0   |
| Best Local Similarity:                              | 100.00%  | Mismatches:   | 0   |
| Query Match:  | 29.64%   | Indels:   | 0   |
| DB:   | 3        | Gaps:   | 0   |
| US-09-270-437D-5 (1-1708) x US-09-261-855-2 (1-577) |          |   |     |
| QY  | 215      | GTGACATCCCCCTTCGGCTCTCGTGCCACCACCATGATGGTGCCATTATTGGCAAG      | 274 |
| Db  | 194      | ValAspfileProLeuArgLeuValProThrGlnTyrValGlyAlaIleileGlyLys    | 213 |
| QY  | 275      | GAGGGGGCCACCATCCGACACATCACAACAGACCCAGTCCCAAGATACACGTGCATAGG   | 334 |
| Db  | 214      | GlucGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg | 233 |

```

QY 335 AAGGAGACGCGAGGTGCGAGCTGAAAGCCCATCAGTGTGCACTCCACCCCTGAGGGTGC 394
Db 234 LysGluAsnAlaGlyAlaAlaGluLysAlaIleSerValHisSerThrProGluGlyCys 253
QY 395 TCCTCCGCTTGAAGATCATCTTGAGATTATGATTAAGAGGCTAAGGACACCAAAACG 454
Db 254 SerSerAlaCysLysMetIleGluGluMetHisLysGluAlaLysAspThrLysThr 273
QY 455 GCTGACGAGGTTCCCTCAAGATCCTGCCCATATAACTTTGTAGGGCGTCTCATTTGGC 514
Db 274 AlaAspGluValProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
QY 515 AAGGAGACGCGAGGCTGAGGAGGTAGCAAGATACCCGAGACAAAATCACCATCTCC 574
Db 294 LysGluGlyArgAsnLeuLysValGluGlnAspThrGluThrLysIleThrIleSer 313
QY 575 TCCTTGCAAGACCTTACCCTTTACAACCTCGAGAGGACCATCACTGTCAAGGGGGCCATC 634
Db 314 SerLeuGlnAspLeuThrLeuTy-AsnProGluArgThrIleThrValLysGlyAlaIle 333
QY 635 GAGATTCTTCAGGGCCGACAGCAATAATGAAGAAAGTTCCGGAGGGCCTATGAGAAAT 694
Db 334 GluAsnCysCysArgAlaGluGlnGluIleMetLysLysValArgGluAlaTyGluAsn 353
QY 695 GATGTGGCTGCATGAGC 712
Db 354 AspValAlaAlaMetSer 359

RESULT 2
US-09-261-855-18
; Sequence 18, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-261-855-18

Alignment Scores:
Pred. No.: 8,95e-40 Length: 49
Score: 49.00 Matches: 49
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.75% Indels: 0
DB: 3 Gaps: 0

US-09-270-437D-5 (1-1708) x US-09-261-855-18 (1-49)
QY 476 ATCTGCGCCCATATAACTTTGTAGGCGTCTCATTTGGCAAGGAGCGAAGCACTGAAG 535
Db 1 IleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLys 20
QY 536 AAGGTAGACGACGATACCGACACAAAATCACCATCTCCCTCGTTGCAAGACCTTACCCTT 595
Db 21 LysValGluGlnAspThrGluThrLysIleThrIleSerSerLeuGlnAspLeuThrLeu 40
QY 596 TACAACCTCGAGAGCACTCATCTGTG 622
Db 41 TyrAsnProGluArgThrIleThrVal 49

RESULT 3
US-09-261-855-20
; Sequence 20, Application US/09261855A

```

```

; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-261-855-20

Alignment Scores:
Pred. No.: 8,8e-39 Length: 48
Score: 48.00 Matches: 48
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.57% Indels: 0
DB: 3 Gaps: 0

US-09-270-437D-5 (1-1708) x US-09-261-855-20 (1-48)
QY 1103 ATACGTGTGCCAGCATCAGCAGCTGGCCGGTCAATTGGCAAGGTGGAAAAACGGTGAAC 1162
Db 1 IleArgValProAlaSerAlaAlaGlyArgValIleGlyLysGlyLysThrValAsn 20
QY 1163 GAGTTGCAGAAATTCACCGGCAGCTGAGGTGGTGTAGTACCAAGAGACCAGACCCCTGATGAG 1222
Db 21 GluLeuGlnAsnLeuThrAlaAlaGluValValProArgAspGlnThrProAspGlu 40
QY 1223 AACGACCAAGTCATCTGTAATAATC 1246
Db 41 AsnAspGlnValIleValLysIle 48

RESULT 4
US-09-261-855-17
; Sequence 17, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-261-855-17

Alignment Scores:
Pred. No.: 8,65e-38 Length: 47
Score: 47.00 Matches: 47
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.39% Indels: 0
DB: 3 Gaps: 0

US-09-270-437D-5 (1-1708) x US-09-261-855-17 (1-47)
QY 233 CTCCTGTGTGCCACCCAGTATGTGGTGCATTTATTCGCAAGGAGGGGCGCACCATCCG 292
Db 1 LeuLeuValProThrGlnTy-ValIleValIleGlyLysGluGlyAlaThrIleArg 20
QY 293 AACATCAAAACAGACCCAGTCCAGATAGACGTGCATAGGAGAGAGAACGCGGTGCA 352

```

```

; ORGANISM: Homo sapiens
US-09-643-597-176

Alignment Scores:
Pred. No.: 5,28e-27 Length: 579
Score: 36.00 Matches: 36
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.43% Indels: 0
DB: 4 Gaps: 0

US-09-270-437D-5 (1-1708) x US-09-643-597-176 (1-579)
QY 254 GTGGGTGCATTATTGGCAAGAGGGGCGCCACCATCGCAACATCACAAAACAGACCCAG 313
Db 207 VAIGLYAlaIleIGLYGSLUGLYAlaThrIleArgAsnIleThrLysGlnThrGln 226
QY 314 TCCAAGATAGACGTGCATAGGAAGAGAAACGACGAGTGCAGCTGAAAAA 361
Db 227 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 242

RESULT 7
US-09-643-597-348
; Sequence 348, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 348
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-643-597-348

Alignment Scores:
Pred. No.: 5,28e-27 Length: 579
Score: 36.00 Matches: 36
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.43% Indels: 0
DB: 4 Gaps: 0

US-09-270-437D-5 (1-1708) x US-09-643-597-348 (1-579)
QY 254 GTGGGTGCATTATTGGCAAGAGGGGCGCCACCATCGCAACATCACAAAACAGACCCAG 313
Db 207 VAIGLYAlaIleIGLYGSLUGLYAlaThrIleArgAsnIleThrLysGlnThrGln 226
QY 314 TCCAAGATAGACGTGCATAGGAAGAGAAACGACGAGTGCAGCTGAAAAA 361
Db 227 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 242

RESULT 8
US-09-480-884A-176
; Sequence 176, Application US/09480884A
; Patent No. 6492597
; GENERAL INFORMATION:

```

; APPLICANT: Wang, Tongtong  
 ; APPLICANT: Fan, Liqun  
 ; APPLICANT: Hosken, Nancy A.  
 ; APPLICANT: Kalos, Michael D.  
 ; APPLICANT: Fanger, Gary R.  
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY  
 ; FILE REFERENCE: 210121.455C6  
 ; CURRENT APPLICATION NUMBER: US/09/480.884A  
 ; CURRENT FILING DATE: 2001-08-27  
 ; NUMBER OF SEQ ID NOS: 330  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 176  
 ; LENGTH: 579  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-480-884A-176

Alignment Scores:  
 Pred. No.: 5,28e-27 Length: 579  
 Score: 36.00 Matches: 36  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 6.43% Indels: 0  
 DB: 4 Gaps: 0

US-09-270-437D-5 (1-1708) x US-09-480-884A-176 (1-579)

QY 254 GTGGTGCCATTATGGCAAGAGGGGCCACCATCCGCAACATCACAAAACAGACCAG 313  
 Db 207 ValGlyAlaIleIleGlyLeuGluGlyAlaThrIleArgAsnIleThrLysGlnThrGln 226  
 QY 314 TCCAGATAGACGTGCATAGGAAGAGAACCGACGTGCGAGTGAAGAAA 361  
 Db 227 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaGluLys 242

## RESULT 9

US-09-542-615A-176

; Sequence 176, Application US/09542615A

; Patent No. 6518256

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; APPLICANT: Fan, Liqun

; APPLICANT: Kalos, Michael D.

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: Hosken, Nancy A.

; APPLICANT: Fanger, Gary R.

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY

; FILE REFERENCE: 210121.455C6

; CURRENT APPLICATION NUMBER: US/09/542.615A

; CURRENT FILING DATE: 2000-04-14

; NUMBER OF SEQ ID NOS: 350

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 176

; LENGTH: 579

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-542-615A-176

Alignment Scores:  
 Pred. No.: 5,28e-27 Length: 579  
 Score: 36.00 Matches: 36  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 6.43% Indels: 0  
 DB: 4 Gaps: 0

US-09-270-437D-5 (1-1708) x US-09-542-615A-176 (1-579)

QY 254 GTGGTGCCATTATGGCAAGAGGGGCCACCATCCGCAACATCACAAAACAGACCAG 313  
 Db 207 ValGlyAlaIleIleGlyLeuGluGlyAlaThrIleArgAsnIleThrLysGlnThrGln 226

QY 314 TCCAGATAGACGTGCATAGGAAGAGAACCGACGTGCGAGTGAAGAAA 361  
 Db 227 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaGluLys 242

## RESULT 10

US-09-542-615A-348

; Sequence 348, Application US/09542615A

; Patent No. 6518256

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; APPLICANT: Fan, Liqun

; APPLICANT: Kalos, Michael D.

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: Hosken, Nancy A.

; APPLICANT: Fanger, Gary R.

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY

; FILE REFERENCE: 210121.455C8

; CURRENT APPLICATION NUMBER: US/09/542.615A

; CURRENT FILING DATE: 2000-04-14

; NUMBER OF SEQ ID NOS: 350

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 348

; LENGTH: 579

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-542-615A-348

Alignment Scores:  
 Pred. No.: 5,28e-27 Length: 579  
 Score: 36.00 Matches: 36  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 6.43% Indels: 0  
 DB: 4 Gaps: 0

US-09-270-437D-5 (1-1708) x US-09-542-615A-348 (1-579)

QY 254 GTGGTGCCATTATGGCAAGAGGGGCCACCATCCGCAACATCACAAAACAGACCAG 313  
 Db 207 ValGlyAlaIleIleGlyLeuGluGlyAlaThrIleArgAsnIleThrLysGlnThrGln 226

QY 314 TCCAGATAGACGTGCATAGGAAGAGAACCGACGTGCGAGTGAAGAAA 361  
 Db 227 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaGluLys 242

## RESULT 11

US-09-606-421B-176

; Sequence 176, Application US/09606421B

; Patent No. 6531315

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; APPLICANT: Fan, Liqun

; APPLICANT: Kalos, Michael D.

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: Hosken, Nancy

; APPLICANT: Fanger, Gary R.

; APPLICANT: Li, Samuel X.

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: 210121.455C9

; CURRENT APPLICATION NUMBER: US/09/606.421B

; CURRENT FILING DATE: 2000-06-28

; NUMBER OF SEQ ID NOS: 358

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 176

; LENGTH: 579

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-606-421B-176

Alignment Scores:  
Pred. No.: 5,28e-27 Length: 579  
Score: 36.00 Matches: 36  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.43% Indels: 0  
DB: 4 Gaps: 0

US-09-270-437D-5 (1-1708) x US-09-606-421B-176 (1-579)

QY 254 GTGGTCCCATTTTGGCAAGAGGGGGCCACCATCCGCAACATCCAAACACAGACCCAG 313  
Db 207 ValGlyAlaIlelleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGln 226  
QY 314 TCCAAGATAGACGTGCATAGGAAGAGAGACGCGAGTGCAGCTGAAAAA 361  
Db 227 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 242

#### RESULT 12

US-09-606-421B-348  
; Sequence 348, Application US/09606421B  
; Patent No. 6531315

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; APPLICANT: Fan, Liqun

; APPLICANT: Kalos, Michael D.

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: Hosken, Nancy

; APPLICANT: Fanger, Gary R.

; APPLICANT: Li, Samuel X.

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER

; FILE REFERENCE: 210121.455C9

; CURRENT APPLICATION NUMBER: US/09/606,421B

; CURRENT FILING DATE: 2000-06-28

; NUMBER OF SEQ ID NOS: 358

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 348

; LENGTH: 579

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-606-421B-348

Alignment Scores:  
Pred. No.: 5,28e-27 Length: 579  
Score: 36.00 Matches: 36  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.43% Indels: 0  
DB: 4 Gaps: 0

US-09-270-437D-5 (1-1708) x US-09-606-421B-348 (1-579)

QY 254 GTGGTCCCATTTTGGCAAGAGGGGGCCACCATCCGCAACATCCAAACACAGACCCAG 313  
Db 207 ValGlyAlaIlelleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGln 226  
QY 314 TCCAAGATAGACGTGCATAGGAAGAGAGACGCGAGTGCAGCTGAAAAA 361  
Db 227 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 242

#### RESULT 13

US-09-261-855-21

; Sequence 21, Application US/09261855A

; Patent No. 6255055

; GENERAL INFORMATION:

; APPLICANT: Ross, Jeffrey

; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN

; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE

; FILE REFERENCE: 960296.95131

; CURRENT APPLICATION NUMBER: US/09/261,855A  
; CURRENT FILING DATE: 1999-03-03  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 21  
; LENGTH: 47  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-261-855-21

#### Alignment Scores:

Pred. No.: 6.01e-19 Length: 47  
Score: 28.00 Matches: 28  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.00% Indels: 0  
DB: 3 Gaps: 0

US-09-270-437D-5 (1-1708) x US-09-261-855-21 (1-47)

QY 278 GGGGCCCATCCGCAACATCCAAACACAGACCCAGTCCAGATAGACGTGCATAGGAAG 337  
Db 16 GlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLys 35  
QY 338 GAGAACGCGAGTGCAGCTGAAAAA 361  
Db 36 GluAsnAlaGlyAlaAlaGluLys 43

#### RESULT 14

US-09-261-855-24

; Sequence 24, Application US/09261855A

; Patent No. 6255055

; GENERAL INFORMATION:

; APPLICANT: Ross, Jeffrey

; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN

; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE

; FILE REFERENCE: 960296.95131

; CURRENT APPLICATION NUMBER: US/09/261,855A

; CURRENT FILING DATE: 1999-03-03

; NUMBER OF SEQ ID NOS: 46

; SOFTWARE: Patent In Ver. 2.0

; SEQ ID NO 24

; LENGTH: 48

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-261-855-24

Alignment Scores:  
Pred. No.: 5.04e-10 Length: 48  
Score: 19.00 Matches: 19  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 3.39% Indels: 0  
DB: 3 Gaps: 0

US-09-270-437D-5 (1-1708) x US-09-261-855-24 (1-48)

QY 1121 GCAGCTGGCGGTCTATGTGCAAAAGTGGAACCGTGAACGAGTTGCAGATTTC 1177  
Db 7 AlaAlaGlyArgValIlelleGlyLysGlyGlyLysThrValAsnGluLeuGlnAsnLeu 25

#### RESULT 15

US-09-261-855-3

; Sequence 3, Application US/09261855A

; Patent No. 6255055

; GENERAL INFORMATION:

; APPLICANT: Ross, Jeffrey

; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN

; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE

; FILE REFERENCE: 960296.95131

; CURRENT APPLICATION NUMBER: US/09/261,855A

; CURRENT FILING DATE: 1999-03-03

; NUMBER OF SEQ ID NOS: 46

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-261-855-3

Alignment Scores:
  Pred. No.:      5.27e-05      Length:      14
  Score:          14.00         Matches:      14
  Percent Similarity: 100.00%   Conservative: 0
  Best Local Similarity: 100.00% Mismatches: 0
  Query Match:      2.50%       Indels: 0
  DB:                3          Gaps: 0

US-09-270-437D-5 (1-1708) x US-09-261-855-3 (1-14)

QY 134 CGCCGAGGGGCTTTGGCTCTCGGGTCTCAGCCGCCAGGGC 175
Db 1 ArgArgGlyGlyPheGlySerArgGlyGlnProArgGingly 14
```

Search completed: July 13, 2004, 12:36:38  
Job time : 28 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: July 13, 2004, 11:51:37 ; Search time 82 Seconds  
(without alignments)  
11770.516 Million cell updates/sec

Title: US-09-270-437D-5  
Perfect score: 3110  
Sequence: 1 agggagcgtgcgcacgcgc.....attctcttcaggttttaaaa 1708

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 3172214

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+n2p-model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_P/US09270437/runat\_13072004\_121921\_9584/app\_query.fasta\_1.1863  
-DB=A Geneseq\_29Jan04 -Qfmt=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -NAXLEN=2000000000  
-USSE=US09270437 @CGN 1 1 154 runat\_13072004\_121921\_9584 -NCPU=6 -ICPU=3  
-NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A Geneseq\_29Jan04.\*  
1: Geneseqp1980s.\*  
2: Geneseqp1990s.\*  
3: Geneseqp2000s.\*  
4: Geneseqp2001s.\*  
5: Geneseqp2002s.\*  
6: Geneseqp2003as.\*  
7: Geneseqp2003bs.\*  
8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID         | Description        |
|------------|-------|-------------|--------|------------|--------------------|
| 1          | 2208  | 71.0        | 577    | 2 AAY30649 | Aay30649 A murine  |
| 2          | 1639  | 52.7        | 579    | 3 AAB11365 | Aab11365 Human lun |
| 3          | 1639  | 52.7        | 579    | 5 ABB75053 | Abb75053 Human lun |
| 4          | 1639  | 52.7        | 579    | 5 ABB74997 | Abb74997 Human lun |
| 5          | 1639  | 52.7        | 579    | 5 ABB75054 | Abb75054 Human lun |
| 6          | 1639  | 52.7        | 579    | 5 ABB61917 | Abp61917 Human lun |
| 7          | 1639  | 52.7        | 579    | 5 ABB61974 | Abp61974 Human lun |
| 8          | 1639  | 52.7        | 579    | 5 ABB61973 | Abp61973 Human lun |
| 9          | 1639  | 52.7        | 579    | 7 ADA28536 | Ada28536 Recombina |
| 10         | 1639  | 52.7        | 579    | 7 ADA28539 | Ada28539 Recombina |

|    |        |      |     |            |                    |
|----|--------|------|-----|------------|--------------------|
| 11 | 1639   | 52.7 | 579 | 7 ADA28438 | Ada28438 Human lun |
| 12 | 1639   | 52.7 | 586 | 5 ABB75048 | Abb75048 Human lun |
| 13 | 1639   | 52.7 | 586 | 5 ABB61968 | Abp61968 Human lun |
| 14 | 1639   | 52.7 | 586 | 7 ADA28517 | Ada28517 Recombina |
| 15 | 1637   | 52.6 | 579 | 3 AAB11328 | Aab11328 Human lun |
| 16 | 1637   | 52.6 | 579 | 3 ABB74960 | Abb74960 Human lun |
| 17 | 1637   | 52.6 | 579 | 5 ABB61880 | Abp61880 Human lun |
| 18 | 1637   | 52.6 | 579 | 7 ADA28266 | Ada28266 Human lun |
| 19 | 1637   | 52.6 | 579 | 7 ADD14066 | Add14066 Human src |
| 20 | 1637   | 52.6 | 579 | 7 ADE53471 | Ade53471 Human lun |
| 21 | 1585   | 51.0 | 319 | 4 AAM93826 | Aam93826 Human pol |
| 22 | 1572   | 50.5 | 619 | 4 ABG21963 | Abg21963 Novel hum |
| 23 | 1530.5 | 49.2 | 614 | 4 ABG06794 | Abg06794 Novel hum |
| 24 | 1530.5 | 49.2 | 620 | 4 AAU16163 | Aau16163 Human nov |
| 25 | 1530.5 | 49.2 | 620 | 6 ABG12592 | Abg12592 Novel hum |
| 26 | 1518   | 48.8 | 583 | 4 ABG12592 | Abg12592 Novel hum |
| 27 | 1464   | 47.1 | 556 | 5 ABG96346 | Abg96346 Human ova |
| 28 | 1464   | 47.1 | 594 | 4 ABG06795 | Abg06795 Novel hum |
| 29 | 1355.5 | 43.6 | 555 | 6 ABU89799 | Abu89799 Novel hum |
| 30 | 1258   | 40.5 | 261 | 4 AAU16161 | Aau16161 Human nov |
| 31 | 1258   | 40.5 | 261 | 6 ABUS5230 | Abu55230 Human nov |
| 32 | 1228   | 39.5 | 250 | 4 AAU16579 | Aau16579 Human nov |
| 33 | 1228   | 39.5 | 250 | 6 ABUS5648 | Abu55648 Human nov |
| 34 | 858    | 27.6 | 558 | 4 ABB58367 | Abb58367 Drosophil |
| 35 | 653    | 21.0 | 171 | 4 AAU16166 | Aau16166 Human nov |
| 36 | 652    | 21.0 | 171 | 6 ABUS5235 | Abu55235 Human nov |
| 37 | 605.5  | 19.5 | 187 | 6 AAU16164 | Aau16164 Human nov |
| 38 | 510    | 16.4 | 192 | 4 AAU16164 | Aau16164 Human nov |
| 39 | 510    | 16.4 | 192 | 6 ABUS5233 | Abu55233 Human nov |
| 40 | 502    | 16.1 | 171 | 4 AAU16583 | Aau16583 Human nov |
| 41 | 502    | 16.1 | 171 | 6 ABUS5652 | Abu55652 Human nov |
| 42 | 465.5  | 15.0 | 148 | 4 ABG21962 | Abg21962 Novel hum |
| 43 | 410    | 13.2 | 266 | 4 ABG12593 | Abg12593 Novel hum |
| 44 | 409.5  | 13.2 | 209 | 4 ABG21961 | Abg21961 Novel hum |
| 45 | 292    | 9.4  | 93  | 4 AAM38501 | Aam38501 Peptide # |

ALIGNMENTS

RESULT 1  
AAY30649  
ID AAY30649 standard; protein; 577 AA.

XX

AC AAY30649;

XX 17-NOV-1999 (first entry)

XX

DE A murine c-myc coding region determinant binding protein.

XX

KW c-myc coding region determinant binding protein; CRD-BP; tumor; c-myc;

KW endonucleolytic attack; half-life; breast cancer; colon cancer;

KW pancreatic cancer.

XX

OS Mus musculus.

XX

PN WO9946594-A2.

XX

PD 16-SEP-1999.

XX

PF 05-MAR-1999; 99WO-US004897.

XX

PR 09-MAR-1998; 98US-0077372P.

XX

XX (WISC ) WISCONSIN ALUMNI RES FOUND.

XX Ross J;

XX WPI; 1999-551506/46.

XX N-P8DB; AAZ10617.

XX Diagnosing presence or absence of a tumor in a human by examining c-myc coding region determinant-binding protein.

PS Example; Fig 1A-D; 79pp; English.

XX The present sequence represents a murine c-myc coding region determinant  
CC binding protein (CRP-BP). The presence or absence of a tumor can be  
CC determined by determining the levels of CRP-BP present in the suspect  
CC tissue, where the CRP-BP shields c-myc RNA from endonucleolytic attack  
CC and so prolongs its half-life. The methods are used for diagnosing  
CC presence or absence of a tumor in a human, especially breast, colon and  
CC pancreatic cancer. They are also used to inhibit cancer cell growth  
XX Sequence 577 AA;

#### Alignment Scores:

Pred. No.: 9,76e-198 Length: 577  
Score: 2208.00 Matches: 438  
Percent Similarity: 98.65% Conservative: 1  
Best Local Similarity: 98.43% Mismatches: 4  
Query Match: 71.00% Indels: 2  
DB: 2 Gaps: 1

US-09-270-437D-5 (1-1708) x AAY30649 (1-577)

```

QY 32 CGGGAGCCATCATGAAGCTGAATGGCCACAGTTGGAGAACCATGCCCTGAAGGCTCC 91
DB 133 ArgGlnAlaIleMetLysLeuAsnGlyHisGlnLeuGluAsnHisAlaLeuLysValSer 152
QY 92 TACATCCCGATGAGCAGATAGCACAGGACCTGAGAAATGGCGCCGAGGGGCTTTGGC 151
DB 153 TyrIleProAspGluGlnIleThrGlnGlyProGluAsnGlyArgArgGlyGlyPheGly 172
QY 152 TCTCGGGTACGCCCGCCAGGGCTCACCTGTGGCAGCGGGGCCCGCCAGCCAGCAGCAG 211
DB 173 SerArgGlyGlnProArgGlnGlySerProValAlaAlaGlyAlaProAlaLysGlnGln 192
QY 212 CAAGTGACATCCCGCTTCGGCTCTCGTCCGCCACCCAGATGTGGGTGCCATTATGGC 271
DB 193 ProValAspIleProLeuAsnGlyLeuLeuValProThrGlnTyrValGlyAlaIleGly 212
QY 272 AAGGAGGGGCCCATCCGACATCCGACATCACAAACAGACCCAGTCCAGATAGCTGCAT 331
DB 213 LysGluGlyAlaThrIleAsnGlnIleThrLysGlnThrGlnSerLysIleAspValHis 232
QY 332 AGGAAGGAGAACGCGAGGTGAGCTGAAAGGCCATCAGTGTGCATCCACCCCTGAGGGC 391
DB 233 ArgLysGluAsnAlaGlyAlaAlaGluLysAlaIleSerValHisSerThrProGluGly 252
QY 392 TGCTCCCTCCGCTGTAGATGATCTGGAGATTAATGATTAAGAGCTAAGGACACAA 451
DB 253 CysSerAlaCysLysMetIleLeuGluIleMetHisLysGluAlaLysAspThrLys 272
QY 452 ACGGCTGACGAGGTCCCTGAGATCCTGGCCCATTAATTAATTTGTAGGGCTCTCATT 511
DB 273 ThrAlaAspGluValProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIle 292
QY 512 GGCAAGGAGACCGAACCTGAGAGAGGTAGACAGATACCGAGACAAATAATCACCATC 571
DB 293 GlyLysGluGlyArgAsnLeuLysLysValGluGlnAspThrGluThrLysIleThrIle 312
QY 572 TCCTCGTTGCAAGACCTTACCTTTACACCTGAGAGGACCATCCTGTGAAGGGGGCC 631
DB 313 SerSerLeuGlnAspLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAla 332
QY 632 ATCGAGATTTGTCAGGGCCGAGAGCAATTAATGAAGAACTTCGGAGGCTTATGAG 691
DB 333 IleGluAsnCysCysArgAlaGluGlnIleMetLysLysValArgGluAlaTyrGlu 352
QY 692 AATGATCTGCTGCATGAGC-----TCTCACCTGATCCCTGGCTGACCTGGCTGCT 745
DB 353 AsnAspValAlaAlaMetSerLeuGlnSerHisLeuIleProGlyLeuAsnLeuAlaAla 372
QY 746 GTAGGTCTTTTCCAGGCTTATCCAGCGAGTCCCGCGCTCCCGAGCAGCGTTACTGGG 805
DB 373 ValGlyLeuPheProAlaSerSerAlaValProProProProSerValThrGly 392

```

```

QY 806 GTGTCTCCTATAGCTCCTTTATGCGAGCTCCCGAGCAGAGATGGTGCAGGTCTTATC 865
DB 393 AlaAlaProTyrSerSerPheMetGlnAlaProGluGlnGluMetValGlnValPheIle 412
QY 866 CCCGCCAGGAGGAGGGCCCATCATCGGCAAGAGGGGAGGACATCAACACGCTCTCC 925
DB 413 ProAlaGlnAlaValGlyAlaIleIleGlyLysGlyGlnHisIleLysGlnLeuSer 432
QY 926 CGTTTTCAGCGCTCCATCAAGATTGCACACCCGAAACACCTGACTCCAAAGTTTCGT 985
DB 433 ArgPheAlaSerAlaSerIleLysIleAlaProProGluThrProAspSerLysValArg 452
QY 986 ATGTTATCATCATCTGGACCCCGAGGCCCAATCAAGCTCAGGGAAGAAATCTATGGC 1045
DB 453 MetValValIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrGly 472
QY 1046 AAACCTCAAGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACCCACATA 1105
DB 473 LysLeuLysGluGluAsnPhePheGlyProLysGluGluValLysLeuGluThrHisIle 492
QY 1106 CGTGTCCAGCATCAGCAGCTGGCGGGTCAATTGGCAAGGTGGAACCGGTGAACGAG 1165
DB 493 ArgValProAlaSerAlaAlaGlyArgValIleGlyLysGlyLysThrValAsnGlu 512
QY 1166 TTGCGAATTTGACGGCAGCTGAGTGGTAGTACCAAGAGACCAAGCCCTGATGAGAAC 1225
DB 513 LeuGlnAsnLeuThrAlaAlaGluValValProArgAspGlnThrProAspGluAsn 532
QY 1226 GACGAGGTCACTCGTCAAAATCATCGACATTTCTATGCCAGTCCAGTGGTCAACGGAAG 1285
DB 533 AspGlnValIleValLysIleIleGlyHisPheTyrAlaSerGlnMetAlaGlnArgLys 552
QY 1286 ATCCGAGACATCTCGGCCAGGTTAAGCAGCAGCATCAGAAGGACAGAGTAACCGAGCC 1345
DB 553 IleArgAspIleLeuAlaGlnValLysGlnGlnHisGlnLysGlyGlnSerAsnLeuAla 572
QY 1346 CAGGACCGGAGGAAG 1360
DB 573 GlnAlaArgArgLys 577

```

#### RESULT 2

```

AAB11365
ID AAB11365 standard; protein; 579 AA.
XX AAB11365;
AC AAB11365;
XX 21-FEB-2001 (first entry)
DT Human lung cancer associated antigen L523S.
DE Lung cancer; therapy; treatment; human; tumor; immunogenic; cytostatic;
KW vaccine; detection.
XX Homo sapiens.
OS WO200061612-A2.
XX 19-OCT-2000.
PD 03-APR-2000; 2000WO-US008896.
XX 02-APR-1999; 99US-00285479.
PR 17-DEC-1999; 99US-00466396.
PR 30-DEC-1999; 99US-00476496.
PR 10-JAN-2000; 2000US-00480884.
PR 22-FEB-2000; 2000US-00510376.
XX (CORI-) CORIYA CORP.
XX Wang T, Fan L;
XX WPI; 2000-628399/60.
DR

```



DR N-PSDB; AAC66035.

XX Isolated polypeptide comprising an immunogenic portion of a lung tumor  
 PT protein is used for detecting and monitoring progression of lung cancer  
 PT in a patient.

XX Claim 3; Page 259-261; 261pp; English.

XX This invention describes a novel isolated polypeptide (I) which  
 CC comprising an immunogenic portion of a lung tumor protein or variant (P2)  
 CC which have cytostatic activity. The polypeptides and polynucleotides are  
 CC used in compositions and vaccines to inhibit the development of cancer,  
 CC especially lung cancer, in a patient. Methods described in the invention  
 CC can be used to monitor the progression of a cancer by carrying out the  
 CC detection at subsequent time points and comparing the results from the  
 CC different time points. CD4+ and/or CD8+ T-Cells isolated from a patient  
 CC are treated with P2, polynucleotides encoding P2 or antigen presenting  
 CC cells expressing P2 and then administered to the patient to inhibit  
 CC development of cancer

XX Sequence 579 AA;

## Alignment Scores:

Pred. No.: 2,36e-144 Length: 579  
 Score: 1639.00 Matches: 336  
 Percent Similarity: 83.89% Conservative: 44  
 Best Local Similarity: 74.17% Mismatches: 57  
 Query Match: 52.70% Indels: 16  
 DB: Gaps: 9

US-09-270-437D-5 (1-1708) x AAB11365 (1-579)

Qy 32 CGGGAGCCATCATGAAGTGAATGGCCACAGTGGAGAACCATGCCCTGAAGGTCTCC 91  
 Db 133 ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuValAla 152  
 Qy 92 TACATCCCGATGACGATAGCA---CAGGGACCTGAGATGGGCCGCCAGGG----- 142  
 Db 153 TyrIleProAspGluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArg 172  
 Qy 143 GCCTTTGGCTCTCGGGGTGACGCCCGCCAGGGCTCACCTGTGGCAGCGGGGCCACGCC 202  
 Db 173 GlyLeuGlyGlnArgGlySerSerArgGlnGlySerPro-----GlySerValSer 189  
 Qy 203 AAGCAGCAGCAAGTGGACATCCCTTCGGCTCTGGTGTGCCACCCAGTATGGGTGCC 262  
 Db 190 LysGlnLysProCysAspLeuProLeuArgLeuValProThrGlnPheValGlyAla 209  
 Qy 263 ATTATTGGCAGAGGGGGCCACCATCCGCAACATCAAAACAGACCCAGTCCAGATA 322  
 Db 210 IleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIle 229  
 Qy 323 GACGTGATAGGAAGGAGACGCGAGGTGCGAGCTGAAAAGCCATCGTGTGCACTCCACC 382  
 Db 230 AspValHisArgLysGluAsnAlaGlyAlaGluLysSerIleThrIleLeuSerThr 249  
 Qy 383 CCTGAGGGCTCCTCGCTGTTGTAAGATGATCTTGGAGATTATGCAATAAGAGGCTAAG 442  
 Db 250 ProGluGlyThrSerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGln 269  
 Qy 443 GACACCAAACGGCTGACGAGGTTCCTGTGAAGATCCTGGCCCAATAAATCTTGTAGGG 502  
 Db 270 AspileLysPheThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGly 289  
 Qy 503 CGTCTCATTTGCAAGGAGCGAAGCTGTAAGAGGTAGAGCAAGATACCGAGACAAA 562  
 Db 290 ArgLeuIleGlyLysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLys 309  
 Qy 563 ATCCACCATCTCCTCGTTCGACAGACTTACCTTTACAACTGAGAGCACCATCTGTG 622  
 Db 310 IleThrIleSerProLeuGlnGlnLeuThrLeuTyrAsnProGluArgThrIleThrVal 329  
 Qy 623 AAGGGGGCCATCGAAGATTGTTCAGGGCCGAGCAGGAAATATGAAGAAGTTCGGGAG 682

Db 330 LysGlyAsnValGluThrCysAlaLysAlaGluGluGluIleMetLysLysIleArgGlu 349  
 Qy 683 GCGTATGAGAAATGATGTGGCTCCCATGAGC-----TCTCACTGATCCCTGCGCTGAAC 736  
 Db 350 SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn 369  
 Qy 737 CTGGCTGTGTAGGTCTTTTCCAGCTTCATCCAGCGCAGTCCGCGC-----CCT 787  
 Db 370 LeuAsnAlaLeuGlyLeuPheProThrSerGlyMetProProProThrSerGlyPro 389  
 Qy 788 CCACAGCAGGTACTGCGGCTCTCTCCCTATAGCTCTTTATGACAGGCTCCCGAGCAGAG 847  
 Db 390 ProSerAlaMetThr-----ProProTyrProGlnPheGluGlnSer---GluThrGlu 406  
 Qy 848 ATGTGTGAGGTCTTTATCCCGCCAGCGAGTGGCGGCATCATCGGCAAGAGGGGCGAG 907  
 Db 407 ThrValHisLeuPheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGln 426  
 Qy 908 CACATCAAAACAGCTCTCCCGGTTTCCAGCGCTCCATCAAGATTGCAACCCCGAAACA 967  
 Db 427 HisIleLysGlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAla 446  
 Qy 968 CTGACTCCAAAGTTCGTATGTTATCATCTGACCGCCAGAGGCGCAATTCAAGGCT 1027  
 Db 447 ProAspAlaLysValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAla 466  
 Qy 1028 CAGGGAAGAACTATGGCAACTCAAGGAGGAGAACTCTTTGGTCCCAAGGAGGAAGTG 1087  
 Db 467 GlnGlyArgIleTyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluVal 486  
 Qy 1088 AAGCTGGAGACCCACATAGCTGTGCCAGCATCAGCAGCTGGCGGGTTCATTGGCAAGGT 1147  
 Db 487 LysLeuGluAlaHisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGly 506  
 Qy 1148 GCAAAACCGTGAACAGTTCGAGAAATTTGACGGCAGCTGAGTGGTAGTACCAAGAGNC 1207  
 Db 507 GlyLysThrValAsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAsp 526  
 Qy 1208 CAGACCCCTGATGAGAACGACCGAGTCTATCGTGAATAATCATCGGACATTTCTATGCCAGT 1267  
 Db 527 GlnThrProAspGluAsnAspGlnValValValLysIleThrGlyHisPheTyrAlaCys 546  
 Qy 1268 CAGATGGCTCAACGGAAGATCCGAGACATCTTGGCCCGAGGTTAAGCAG---CAGCATCAG 1324  
 Db 547 GlnValAlaGlnArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGln 566  
 Qy 1325 AAGGGA---CAGAGTAACCCAGCGCCAGCGAGGAG 1360  
 Db 567 LysAlaLeuGlnSerGlyProProGlnSerArgArgLys 579  
 RESULT 3  
 ABB75053  
 ID ABB75053 standard; protein; 579 AA.  
 XX  
 AC ABB75053;  
 XX  
 DT 01-MAY-2002 (first entry)  
 XX  
 DE Human lung tumour L52S3 recombinant protein sequence SEQ ID NO:446.  
 XX  
 KW Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;  
 KW immune response.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200200174-A2.  
 XX  
 PD 03-JAN-2002.  
 XX  
 PF 28-JUN-2001; 2001WO-US021065.  
 XX  
 PR 28-JUN-2000; 2000US-00606421.

PR 02-AUG-2000; 2000US-00630940.  
 PR 21-AUG-2000; 2000US-00643597.  
 PR 15-SEP-2000; 2000US-00662786.  
 PR 09-OCT-2000; 2000US-00685596.  
 PR 12-DEC-2000; 2000US-00735705.  
 PR 07-MAY-2001; 2001US-00850716.  
 XX (CORI-) CORIXA CORP.  
 XX  
 PI Wang T, Wang A, Steiky YAW, Li SX, Kalos WD, Henderson RA;  
 PI McNeill PD, Fanger N, Retter MW, Marnerakis M, Fanger GR;  
 PI Vedvick TS, Carter D, Watanabe Y, Peckham DW;  
 XX  
 DR WPI; 2002-090513/12.  
 DR N-PSDB; ABL49297.  
 XX  
 PT Polynucleotides encoding lung tumor polypeptides, useful for treating  
 PT lung cancer or stimulating an immune response.  
 XX  
 PS Claim 2; Page 365-367; 374pp; English.  
 XX  
 CC The present invention describes human lung tumour proteins. Human lung  
 CC tumour proteins and polynucleotides have cytostatic and immunostimulant  
 CC activities, and can be used in vaccine production. Compositions  
 CC comprising the lung tumour proteins, polynucleotides, antibodies, fusion  
 CC proteins, T cell populations, or antigen presenting cells that express  
 CC the lung tumour proteins are useful for treating lung cancer or  
 CC stimulating an immune response. ABL48959 to ABL49300 and ABB74946 to  
 CC ABB75070 represent sequences used in the exemplification of the present  
 CC invention  
 XX  
 SQ Sequence 579 AA;  
 Alignment Scores:  
 Pred. No.: 2,36e-144 Length: 579  
 Score: 1639.00 Matches: 336  
 Percent Similarity: 83.89% Conservative: 44  
 Best Local Similarity: 74.17% Mismatches: 57  
 Query Match: 52.70% Indels: 16  
 DB: 5 Gaps: 9  
 US-09-270-437D-5 (1-1708) x ABB75053 (1-579)  
 QY 32 CGGGAGCATCATGAGCTGAATGGCCACAGTTGGAGAACCATGCCCTGAAGTCTCC 91  
 Db 133 ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla 152  
 QY 92 TACATCCCGCATGACAGATAGCA---CAGGGACCTGAGAAATGGCGCCGAGGG----- 142  
 Db 153 TyrIleProAspGluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyA-Arg 172  
 QY 143 GCGTTTGGCTCTCGGGTACGCCGCCAGGCTCAGCTGTGGCAGGGGGGCCAGCC 202  
 Db 173 GlyLeuGlyGlnArgGlySerArgGlnGlySerPro-----GlySerValSer 189  
 QY 203 AAGCAGCAGCAAGTGGACATCCCTTCGCTCGTCCGCCACCCAGTAGTGGGTGCC 262  
 Db 190 LysGlnLysProCysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAla 209  
 QY 263 ATTATTGGCAGAGGGGGCCACCATCCGCAACATCAAAAACAGACCCAGTCCAGATA 322  
 Db 210 IleIleGlyLysGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIle 229  
 QY 323 GACGTGATAGGAGGAGAGCGAGTGCAGCTGAAAGCCCATCAGTGTGCATCCACC 382  
 Db 230 AspValHisArgLysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThr 249  
 QY 383 CCTGAGGCTCTCTCCGCTTGTAGATGATCTTGGAGATTATGCAATAAGAGGCTAAG 442  
 Db 250 ProGluGlyThrSerAlaAlaCysLysSerIleLeuGluIleMethHisLysGluAlaGln 269  
 QY 443 GACACCAAAACGGCTGACGAGTTCCTCCCTGAAGATCCTGGCCCATTAATCTTTGAGG 502

Db 270 AspIleLysPheThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGly 289  
 QY 503 CGTCTCATTTGGCAAGAGGACGGAACCTGTAAGAAGGTAGAGCAAGATACCGAGACAAA 562  
 Db 290 ArgLeuIleGlyLysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLys 309  
 QY 563 ATCACCATCTCTCTCGTTGCAAGACCTTACACCTTTACAAACCTCGAGAGGACCATCACTGTG 622  
 Db 310 IleThrIleSerProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrVal 329  
 QY 623 AAGGGGCCCATCGAAGATTGTGAGGCCGAGGAGAAATAATCAAGAAATTCGGGAG 682  
 Db 330 LysGlyAsnValGluThrCysAlaLysAlaGluGluIleMetLysLysLysLysLys 349  
 QY 683 GCCTATGAGATGATGTGCTGCCATGAGC-----TCTCACCTGATCCCTGGCTGAAC 736  
 Db 350 SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn 369  
 QY 737 CTGGCTGTGTAGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCCG-----CCT 787  
 Db 370 LeuAsnAlaLeuGlyLeuPheProThrSerGlyMetProProThrSerGlyPro 389  
 QY 788 CCACGACGGTACTGGGGCTGCTCTTATAGCTTATGTCAGGCTCCGAGCAGGAG 847  
 Db 390 ProSerAlaMetThr-----ProProTyrProGlnPheGluGlnSer---GluThrGlu 406  
 QY 848 ATGGTGCAGGTGTTTATCCCGCCGAGCAGTGGCGCCATCATCGGCAAGAGGGGCGAG 907  
 Db 407 ThrValHisLeuPheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGln 426  
 QY 908 CACATCAACAGCTCTCCCGGTTGCGAGCCCTCCATCAAGATGTCACACCCGAAACA 967  
 Db 427 HisIleLysGlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAla 446  
 QY 968 CCTGACTCCAAAGTTGATGTTTATCATCTGACCGCCGAGAGGCCCAATTCAAGGCT 1027  
 Db 447 ProAspAlaLysValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAla 466  
 QY 1028 CAGGGAAGAATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGTTCCCAAGGAGAGTG 1087  
 Db 467 GlnGlyArgIleTyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluVal 486  
 QY 1088 AAGCTGGAGACCCCATACAGTGTGCGCAGCATCAGCAGTGGCGGCTCATTTGGCAAGGT 1147  
 Db 487 LysLeuGluAlaHisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGly 506  
 QY 1148 GGAATAACGGTGAACGAGTTCAGAAATTTGACGCGCAGCTGAGGTGGTAGTCAAGAGAC 1207  
 Db 507 GlyLysThrValAsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAsp 526  
 QY 1208 CAGACCCCTGATGAGAACGACGAGTCTATCTGAAATCATCGGACATTTCTATGCCAGT 1267  
 Db 527 GlnThrProAspGluAsnAspGlnValValLysIleThrGlyHisPheTyrAlaCys 546  
 QY 1268 CAGATGGCTCAACGGAAGATCCGAGACATCTCGGCCAGGTTAAGCAG-----CAGCATCAG 1324  
 Db 547 GlnValAlaGlnArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGln 566  
 QY 1325 AAGGGA---CAGAGTAACCGAGCCAGGCACGGAGGAG 1360  
 Db 567 LysAlaLeuGlnSerGlyProProGlnSerArgArgLys 579  
 RESULT 4  
 ABB74997  
 ID ABB74997 standard; protein; 579 AA.  
 XX ABB74997;  
 XX 01-MAY-2002 (first entry)  
 DE Human lung tumour L523S protein sequence SEQ ID NO:348.  
 XX Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;  
 KW

KW immune response.

XX Homo sapiens.

OS WO200200174-A2.

PN 03-JAN-2002.

XX 28-JUN-2001; 2001WO-US021065.

XX 28-JUN-2000; 2000US-00606421.

PR 02-AUG-2000; 2000US-00630940.

PR 21-AUG-2000; 2000US-00643597.

PR 15-SEP-2000; 2000US-00662786.

PR 09-OCT-2000; 2000US-00685696.

PR 12-DEC-2000; 2000US-00735705.

PR 07-MAY-2001; 2001US-00850716.

XX (CORI-) CORIXA CORP.

PA Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA,

PI McNeill PD, Fanger N, Retter MW, Warnerakis M, Fanger GR;

PI Vedvick TS, Carter D, Watanabe Y, Peckham DW;

XX WPI; 2002-090513/12.

DR N-PSDB; ABL49254.

XX Polynucleotides encoding lung tumor polypeptides, useful for treating

PT lung cancer or stimulating an immune response.

XX Example 2; Page 330-332; 374pp; English.

XX The present invention describes human lung tumor proteins. Human lung

CC tumor proteins and polynucleotides have cytostatic and immunostimulant

CC activities, and can be used in vaccine production. Compositions

CC comprising the lung tumor proteins, polynucleotides, antibodies, fusion

CC proteins, T cell populations, or antigen presenting cells that express

CC the lung tumor proteins are useful for treating lung cancer or

CC stimulating an immune response. ABL48959 to ABL49300 and ABB74946 to

CC ABB75070 represent sequences used in the exemplification of the present

CC invention

XX Sequence 579 AA;

SQ Alignment Scores:

Pred. No.: 2,36e-144 Length: 579

Score: 1639.00 Matches: 336

Percent Similarity: 83.89% Conservative: 44

Best Local Similarity: 74.17% Mismatches: 57

Query Match: 52.70% Indels: 16

DB: 5 Gaps: 9

US-09-270-437D-5 (1-1708) x ABB74997 (1-579)

QY 32 CGGGAGCCATCATGAGCTGAATGGCCACAGTGGGACCATGCGCTCAGGCTCC 91

DB 133 ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGlnAsnPheThrLeuLysValAla 152

QY 92 TACATCCCGGTGAGCAGATAGCA---CAGGGACCTGAGATGGCGCCGAGG----- 142

DB 153 TyrIleProAspGluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArg 172

QY 143 GCCTTGGCTCTCGGGTCAGCCCGGCGGAGGCTCAGCTGGCAGCGGGGCCCCAGCC 202

DB 173 GlyLeuGlyGlnArgGlySerArgGlnGlySerPro-----GlySerValSer 189

QY 203 AAGCAGCAGCAAGTGGACATCCCTCGGCTCTCGGTGGCCGCCAGTATGTGGGTGCC 262

DB 190 LysGlnLysProCysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAla 209

QY 263 ATTATTGCAAGGAGGGGGCCACCATCGGCAACATCAAAACAGACCCAGTCCAGATA 322

DB 210 IleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIle 229

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

323 GACGTGCATAGGAAGAGAAACCCAGGTGCAGTGAATAAAGCCATCAGTGTGCACTCCACC 382

230 AspValHisArgLysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThr 249

383 CCTGAGGGTGTCTCCCTCGCTTGTAGATGATCTTGAGATATATGAGATATATCAAGAGCGTAAG 442

250 ProGluGlyThrSerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGln 269

443 GACACCAAAACCGCTGACGAGGTTCCTCTGAAGATCCTCGCCCAATAAATCTTGTAGGG 502

270 AspIleLysPheThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGly 289

503 CGTCTCATTTGGCAAGAGAGACCGAACCTGAAGAGGTAGAGCAAGATACCGAGACAAA 562

290 ArgLeuIleGlyLysGluGlyArgAsnLeuLysLysIleGluGlnAspThrAspThrLys 309

563 ATCACCATCTCTCTGCTTGAAGACCTTACCTTTACACCTGAGAGGACCATCACTGTG 622

310 IleThrIleSerProLeuGluGluLeuThrLeuTyrAsnProGluArgThrIleThrVal 329

623 AAGGGGCCATCGAGAAATTTGTCAGGCGCGAGAGGAAATTAATGAAGAAAGTTCCGGAG 682

330 LysGlyAsnValGluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGlu 349

683 GCCTATGAGAATGATGTGCTGCCATGAGC-----TCTCAGCTGATCCCTGCGCTGAGC 736

350 SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn 369

737 CTGGTGTGTAGGTCTTTTCCAGCTTCCAGCGCAGTCCCGCG-----CCT 787

370 LeuAsnAlaLeuGlyLeuPheProThrSerGlyMetProProThrSerGlyPro 389

788 CCAGCAGCGGTACCTGGGCTGCTCCCTATAGCTCCTTTATGAGCGTCCCGAGCAGAG 847

390 ProSerAlaMetThr-----ProProTyrProGlnPheGluGlnSer---GluThrGlu 406

848 ATGGTCCAGGTGTATATCCCGCCAGGAGGCGGCGCATCATCGGCAAGAGGGGAG 907

407 ThrValHisLeuPheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGln 426

908 CACATCAAAACAGCTCTCCCGGTTTGGCAGCGCTTCATCAAGATTGCACACCCGAAACA 967

427 HisIleLysGlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAla 446

968 CTGACTCCAAAGTTCGTATGTTATCATCTGACCGCCAGAGCCCAATTCAGAGCT 1027

447 ProAspAlaLysValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAla 466

1028 CAGGGAAGAATCTATGGCAAACTCAAGGAGAGAACTTCTTTGGTCCCAAGGAGGAAGTG 1087

467 GlnGlyArgIleTyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluVal 486

1088 AGCTGGAGACCCACATAGCTGTGCCAGCATCAGCAGCTGGCGGGTCACTTGCCTCAAGGT 1147

487 LysLeuGluAlaHisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGly 506

1148 GGAAAAACCGTGAACAGATTGAGCGGAGCTTTCAGCGAGCTGAGGTGGTAGTACCAAGAGAC 1207

507 GlyLysThrValAsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAsp 526

1208 CAGACCCCTGATAGAACACGAGCTCATCTGTAATAATCATCGGACATTTCTATCCAGT 1267

527 GlnThrProAspGluAsnAspGlnValValLysIleThrGlyHisPheTyrAlaCys 546

1268 CAGATGGCTCAACGGAGATCCGAGACATCTCGGCCCGAGGTAAAGCAG---CAGCATCAG 1324

547 GlnValAlaGlnArgLysIleGlnGluLeuThrGlnValLysGlnHisGlnGlnGln 566

1325 AAGGGA---CAGAGTAACCCAGGCCCGAGCAGCAGGAG 1360

567 LysAlaLeuGlnSerGlyProProGlnSerArgArgLys 579

RESULT 5  
 ABB75054  
 ID ABB75054 standard; protein; 579 AA.  
 XX  
 AC ABB75054;  
 XX  
 DT 01-MAY-2002 (first entry)  
 XX  
 XX Human lung tumour L523S recombinant protein sequence SEQ ID NO:449.  
 DE  
 XX Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;  
 KW immune response.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200200174-A2.  
 XX  
 PD 03-JAN-2002.  
 XX  
 PF 28-JUN-2001; 2001WO-US021065.  
 XX  
 PR 28-JUN-2000; 2000US-00606421.  
 XX  
 PR 02-AUG-2000; 2000US-00630940.  
 XX  
 PR 21-AUG-2000; 2000US-00643597.  
 XX  
 PR 15-SEP-2000; 2000US-00662786.  
 XX  
 PR 09-OCT-2000; 2000US-00685696.  
 XX  
 PR 12-DEC-2000; 2000US-00735705.  
 XX  
 PR 07-MAY-2001; 2001US-00850716.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Wang T, Wang A, Sheiky YAW, Li SX, Kalos MD, Henderson RA;  
 PI McNeill PD, Fanger N, Retter MW, Marnerakis M, Fanger GR;  
 PI Vedvick TS, Carter D, Watanabe Y, Peckham DW;  
 XX  
 DR WPI; 2002-090513/12.  
 DR N-PSDB; ABL49299.  
 XX  
 XX Polynucleotides encoding lung tumor polypeptides, useful for treating  
 PT lung cancer or stimulating an immune response.  
 XX  
 PS Claim 2; Page 368-369; 374pp; English.  
 XX  
 CC The present invention describes human lung tumour proteins. Human lung  
 CC tumour proteins and polynucleotides have cytostatic and immunostimulant  
 CC activities, and can be used in vaccine production. Compositions  
 CC comprising the lung tumour proteins, polynucleotides, antibodies, fusion  
 CC proteins, T cell populations, or antigen presenting cells that express  
 CC the lung tumour proteins are useful for treating lung cancer or  
 CC stimulating an immune response. ABL48959 to ABL49300 and ABB74946 to  
 CC ABB75070 represent sequences used in the exemplification of the present  
 CC invention  
 XX  
 SQ Sequence 579 AA;  
 Alignment Scores:  
 Pred. No.: 2,36e-144 Length: 579  
 Score: 1639.00 Matches: 336  
 Percent Similarity: 83.89% Conservative: 44  
 Best Local Similarity: 74.17% Mismatches: 57  
 Query Match: 52.70% Indels: 16  
 DB: 5 Gaps: 9  
 US-09-270-437D-5 (1-1708) x ABB75054 (1-579)  
 QY 32 CGGGGAGCCATCATGAGCTGAATGGCCACCATGTTGAGAACCATGCTCCAGGCTCC 91  
 DB 133 ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla 152  
 QY 92 TACATCCCGCATGACGATAGCA---CAGGACCTGAGAAATGGGCCCGAGGG----- 142  
 DB 153 TyrileProAspGluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArg 172

143 GGCCTTTGGCTCTCGGGGTACGCCCGCCAGCGCTCACCTGTGGCAGCGGGGCCCCAGCC 202  
 |||||  
 DB GlyLeuGlyGlnArgGlySerSerArgGlnGlySerPro-----GlySerValSer 189  
 QY 203 AAGCAGCAGCAAGTGGACATCCCTTCGGTCTCGTGTGGCCACCCAGTATGTGGTGCC 262  
 |||||  
 DB LysGlnLysProCysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAla 209  
 QY 263 ATTATTGGCAAGGGGGGCCACCATCGCAACATCAAAAACAGACCCAGTCCCAAGATA 322  
 |||||  
 DB IleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIle 229  
 QY 323 GACGTGCATAGAGGAGAACGCGAGTGCAGCTGAAAAGCCATCAGTGTGCATCCACC 382  
 |||||  
 DB AspValHisArgLysGluAsnAlaGlyAlaGlySerIleThrIleLeuSerThr 249  
 QY 383 CCTGAGGGCTGCTCCTCGCTTGAAGATGATCTTGGAGATTATGCATATAAGAGCTAAG 442  
 |||||  
 DB ProGluGlyThrSerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGln 269  
 QY 443 GACACAAAACGGGTGACGAGGTTCCTGAAGATCCTGGCCCATATAATCTTTGTAGG 502  
 |||||  
 DB AspileLysPheThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGly 289  
 QY 503 CGTCTCATTGGCAAGGAGGACGAACTGAAAGAGGTAGAGCAAGATACCGACACAAA 562  
 |||||  
 DB ArgLeuIleGlyLysGluGlyArgAsnLeuLysLysIleGluGlnAspThrAspThrLys 309  
 QY 563 ATCACCATCTCCTCGTTGCAAGACCTTACCCCTTTACAAACCTGAGAGACCATCACTGTG 622  
 |||||  
 DB IleThrIleSerProLeuGlnGluLeuThrLeuTyAsnProGluArgThrIleThrVal 329  
 QY 623 AAGGGGCCATCGAGAAATGTTGAGGGCCGAGCAGGAATAATAGAGAAATGTCGGAG 682  
 |||||  
 DB LysGlyAsnValGluThrCysAlaLysAlaGluGluGluIleMetLysLysIleArgGlu 349  
 QY 683 GCCTATGAGAAATGATGTGGCTGCCATGAGC-----TCTACCTGATCCCTGGCTGAAC 736  
 |||||  
 DB SerTyGluAsnAspileAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn 369  
 QY 737 CTGGCTGCTGTAGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCG-----CCT 787  
 |||||  
 DB LeuAsnAlaLeuGlyLeuPheProThrSerGlyMetProProThrSerGlyPro 389  
 QY 788 CCACGACGGTTACTGGGGCTGCTCCCTATAGTCCCTTATGACAGCTCCGAGCAGAG 847  
 |||||  
 DB ProSerAlaMetThr-----ProProTyProGlnPheGluGlnSer---GluThrGlu 406  
 QY 848 ATGGTGCAAGTGTATCCCGCCAGGAGTGGGGCCCATCATCGGCAAGAGGGGAG 907  
 |||||  
 DB ThrValHisLeuPheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGln 426  
 QY 908 CACATCAACAGCTCTCCCGGTTGCCAGCGCTCCATCAAGATTGACACCCCGCAACA 967  
 |||||  
 DB HisLysGlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAla 446  
 QY 968 CTGACTCCAAAGTTCGTATGTTATCATCTACTGACCGCCAGAGGGCCCAATCAAGCT 1027  
 |||||  
 DB ProAspAlaLysValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAla 466  
 QY 1028 CAGGGAAGAACTATGGCAAACTCAGGAGGAGACTTCTTTGGTCCCAAGGAGGAGTG 1087  
 |||||  
 DB GlnGlyArgIleTyGlyLysIleLysGluGlnAsnPheValSerProLysGluGluVal 486  
 QY 1088 AAGCTGGAGAGCCCATACGTGTGCCAGCATCAGCAGCTGGCGGGTTCATTGGCAAGGT 1147  
 |||||  
 DB LysLeuGluAlaHisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGly 506  
 QY 1148 GAAAAAAGCGTGAACGAGTTCAGAAATTTGACGGCAGCTAGGTGGTAGTACCAAGAGAC 1207  
 |||||  
 DB GlyTyThrValAsnGluLeuGlnAsnLeuSerAlaGluValValProArgAsp 526  
 QY 1208 CAGACCCCTGATGAGACGACCGAGTTCATCGTGAANAATCATCGGACATTTCTATGCCAGT 1267

Db 527 GlnThrProAspGluAsnAspGlnValValValLysIleThrGlyHisPheTyrAlaCys 546  
 QY 1268 CAGATGGCTCAACGAGATCCGACACATCTCGCCAGGTTAAGCAG---CAGCATCAG 1324  
 Db 547 GlnValAlaGlnArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGln 566

QY 1325 AAGGGA---CAGAGTAACCGCCCGCCAGCAGCGGAGGAG 1360  
 Db 567 LysAlaLeuGlnSerGlyProProGlnSerArgArgLys 579

## RESULT 6

ABP61917  
 ID ABP61917 standard; protein; 579 AA.

XX  
 AC ABP61917;  
 DT 07-OCT-2002 (first entry)

DE Human lung cancer associated protein sequence SEQ ID NO:348.  
 DE  
 KW Human; lung cancer; lung tumour; cytostatic; gene therapy; vaccine.  
 XX

OS Homo sapiens.  
 XX  
 FN WO200247534-A2.  
 PN 20-JUN-2002.

XX 30-NOV-2001; 2001WO-US047576.  
 XX 12-DEC-2000; 2000US-00735705.  
 PR 07-MAY-2001; 2001US-00850716.  
 PR 28-JUN-2001; 2001US-00897778.

XX (CORI-) CORIXA CORP.  
 XX Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;  
 FI McNeill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedwick TS;  
 FI Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;  
 XX WPI; 2002-583465/62.  
 DR N-PSDB; ABQ92440.

XX Novel lung carcinoma polynucleotide sequences and polypeptides encoded by  
 PT the polynucleotides, useful in pharmaceutical compositions such as  
 PT vaccines and as markers to indicate the presence of lung cancer.  
 XX  
 PS Example 2; Page 337-339; 381pp; English.

XX The present invention describes isolated human lung carcinoma  
 CC polynucleotides (I) and polypeptides (II). (I) and (II) have cytostatic  
 CC activity, and can be used in gene therapy and in vaccines. Compositions  
 CC comprising (I) or (II) can be used for stimulating an immune response in  
 CC a patient and for treating lung cancer in a patient. Oligonucleotides of  
 CC (I) can be used for detecting the presence of a cancer in a patient, by  
 CC obtaining a biological sample from the patient, contacting the biological  
 CC sample with the oligonucleotide, detecting in the sample, an amount of  
 CC polynucleotide that hybridises to the oligonucleotide and comparing the  
 CC amount of polynucleotide that hybridises to the oligonucleotide to a  
 CC predetermined cut-off value, and determining the presence of a cancer in  
 CC the patient. (I) and (II) are useful in pharmaceutical compositions, e.g.  
 CC vaccines. (I) is useful as a marker to indicate the presence or absence  
 CC of a cancer such as lung cancer. ABQ92145 to ABQ92486 and ABP61866 to  
 CC ABP61992 represent sequences used in the exemplification of the present  
 CC invention

XX Sequence 579 AA;

## Alignment Scores:

Pred. No.: 2,36e-144 Length: 579  
 Score: 1639.00 Matches: 336  
 Percent Similarity: 83.89% Conservative: 44

Best Local Similarity: 74.17% Mismatches: 57  
 Query Match: 52.70% Indels: 16  
 DB: 5 Gaps: 9  
 US-09-270-437D-5 (1-1708) x ABP61917 (1-579)

QY 32 CGGGGAGCCATCATGAGCTGATGCGCCACACAGTTGGAGAACCATGCTCCCTGAAGTCTCC 91  
 Db 133 ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla 152  
 QY 92 TACATCCCGATGAGCAGATAGCA---CAGGACCTGAGAAATGGGCGCGAGGG----- 142  
 Db 153 TyrIleProAspGluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArg 172  
 QY 143 GGCCTTTGGCTCTCGGGGTGAGCCCGCCAGGGCTCACCTGTGGCAGCGGGGCCCCAGCC 202  
 Db 173 GlyLeuGlyGlnArgGlySerSerArgGlnGlySerPro-----GlySerValSer 189  
 QY 203 AAGCAGCAGCAAGTGGACATCCCTTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 262  
 Db 190 LysGlnLysProCysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAla 209  
 QY 263 ATTATTGGCAAGAGGGGCCACCATCCGCAACATCACAAACAGACCCAGTCCCAAGATA 322  
 Db 210 IleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIle 229  
 QY 323 GAGTGCATAGGAAGAGAACGCGAGTGCAGCTGAGAAAGCCATCAGTGTGCACCTCCACC 382  
 Db 230 AspValHisArgLysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThr 249  
 QY 383 CTGAGGGCTGCTCTCCCTTCTAAGATGATCTTGGAGATTATGATATAAGAGGTAAG 442  
 Db 250 ProGluGlyThrSerAlaAlaCysLysSerIleLeuGluIleMethHisLysGluAlaGln 269  
 QY 443 GACACCAAAACGGCTGACAGGTTCCCTGAAGATCTCGCCCATTAATTAATCTTTAGGG 502  
 Db 270 AspIleLysPheThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGly 289  
 QY 503 CGTCTCATTTGGAAGAGGAGCGAACCTGAAGAGGTAGACAAAGATACCCAGACAAA 562  
 Db 290 ArgLeuIleGlyLysGluGlyArgAsnLeuLysLysIleGluGluAspThrAspThrLys 309  
 QY 563 ATCACCATCTCTCTGTTGCAAGACCTTACCTTTCAACCCCTGAGAGGACCATCATCTGG 622  
 Db 310 IleThrIleSerProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrVal 329  
 QY 623 AAGGGGCGCATCGAGAATTGTTGCGAGGCGCAGGAGGAATAATGAAGAAAGTTCGGGAG 682  
 Db 330 LysGlyAsnValGluThrCysAlaLysAlaGluGluIleMethLysLysIleArgGlu 349  
 QY 683 GCCTATGAGATGATGTGGCTGCCATGAGC-----TCTCACCTGATCCCTGGCTGAAC 736  
 Db 350 SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn 369  
 QY 737 CTGGCTGCTGTAGGCTTTTCCAGCTTCATCCAGCGGAGTCCCGCG-----CCT 787  
 Db 370 LeuAsnAlaLeuGlyLeuPheProThrSerGlyMetProProProThrSerGlyPro 389  
 QY 788 CCAGCAGCGGTTACTCGGGGCTCTCCCTATAGCTCTTTATCGAGGCTCCCGAGCAGGAG 847  
 Db 390 ProSerAlaMetThr-----ProProTyrProGlnPheGluGlnSer---GluThrGlu 406  
 QY 848 ATGGTGCAGGTGTTTATCCCGCCCGCAGGAGTGGGCGGCATCATCGGAAGAGGGGAG 907  
 Db 407 ThrValHisLeuPheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGln 426  
 QY 908 CACATCAAAACAGCTCTCCCGGTTTCCAGCGCTCCATCAAGATGACACCCCAACA 967  
 Db 427 HisIleLysGlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAla 446  
 QY 968 CCTGACTCAAAAGTTCGTATGTTATCATCTGACCGCCAGAGCGCCCAATTCAGGCT 1027  
 Db 447 ProAspAlaLysValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAla 466

|          |   |            |          |          |          |            |          |          |         |              |      |
|----------|---|------------|----------|----------|----------|------------|----------|----------|---------|--------------|------|
| QY       | 1028  | CAGGGAAGAA | TCTATGTC | CAAAACTC | CAAGSAGG | AGAACTTCTT | TGTC     | CCCAAGG  | AGAGTG  | 1087         |      |
|          |   |            |          |          |          |            |          |          |         |              |      |
| DB       | 467   | GlnGlyArg  | IleTyr   | GlyLys   | IleLys   | GluLys     | GluLys   | PheVal   | SerPro  | LysGluVal    | 486  |
|          |   |            |          |          |          |            |          |          |         |              |      |
| QY       | 1088  | AAGCTGGAG  | ACCACCA  | TACATG   | TGTC     | CCAGCATC   | AGCAGCTG | CGCGGCTC | ATGCGCA | AGGT         | 1147 |
|          |   |            |          |          |          |            |          |          |         |              |      |
| DB       | 487   | LysLeuGlu  | AlaHis   | IleLeu   | ArgVal   | ProSer     | PheAla   | AlaGly   | ArgVal  | IleGlyLysGly | 506  |
|          |   |            |          |          |          |            |          |          |         |              |      |
| QY       | 1148  | GGAAAAAC   | GGTGAAC  | GAGTTC   | CAAAATT  | TGACGCG    | CAGCTC   | AGGTGGT  | AGTACCA | AGAGAC       | 1207 |
|          |   |            |          |          |          |            |          |          |         |              |      |
| DB       | 507   | GlyLysThr  | ValAsn   | GluLeu   | GlnAsn   | LeuSer     | SerAla   | GluVal   | ValVal  | ProArgAsp    | 526  |
|          |   |            |          |          |          |            |          |          |         |              |      |
| QY       | 1208  | CAGACCCCT  | GATGAG   | AACGAC   | CCAGTCA  | TGTCGTG    | AAATCAT  | CGGACAT  | TTTCTAT | GCCAGT       | 1267 |
|          |   |            |          |          |          |            |          |          |         |              |      |
| DB       | 527   | GlnThrPro  | AspGlu   | AsnAsp   | GlnVal   | ValVal     | LysIle   | ThrGly   | HisPhe  | TyrAlaCys    | 546  |
|          |   |            |          |          |          |            |          |          |         |              |      |
| QY       | 1268  | CAGATGCTC  | CAACGG   | AGATCC   | GAGACAT  | CTCTG      | CCCCAGG  | TTAAGC   | AG--    | --CAGCATCAG  | 1324 |
|          |   |            |          |          |          |            |          |          |         |              |      |
| DB       | 547   | GlnValAla  | GlnArg   | LysIle   | GlnGlu   | IleLeu     | ThrGln   | ValLys   | GlnHis  | GlnGlnGln    | 566  |
|          |   |            |          |          |          |            |          |          |         |              |      |
| QY       | 1325  | AAAGGA--   | --CAGAGT | AAAC     | AGGCC    | CCAGGC     | ACGAGG   | AAAG     | 1360    |              |      |
|          |   |            |          |          |          |            |          |          |         |              |      |
| DB       | 567   | LysAlaLeu  | GlnSer   | GlyPro   | ProGln   | SerArg     | ArgLys   | 579      |         |              |      |
|          |   |            |          |          |          |            |          |          |         |              |      |
| RESULT 7 |   |            |          |          |          |            |          |          |         |              |      |
| ABP61974 |   |            |          |          |          |            |          |          |         |              |      |
| ID       | ABP61974 standard; protein; 579 AA.                                 |            |          |          |          |            |          |          |         |              |      |
| XX       | ABP61974;   |            |          |          |          |            |          |          |         |              |      |
| XX       | AC  |            |          |          |          |            |          |          |         |              |      |
| XX       | DT  |            |          |          |          |            |          |          |         |              |      |
| XX       | 07-OCT-2002 (first entry)   |            |          |          |          |            |          |          |         |              |      |
| DE       | Human lung cancer associated protein sequence SEQ ID NO:449.        |            |          |          |          |            |          |          |         |              |      |
| XX       | Human; lung cancer; lung tumour; cytostatic; gene therapy; vaccine. |            |          |          |          |            |          |          |         |              |      |
| KW       | Homo sapiens.   |            |          |          |          |            |          |          |         |              |      |
| OS       | Homo sapiens.   |            |          |          |          |            |          |          |         |              |      |
| XX       | WC200247534-A2.   |            |          |          |          |            |          |          |         |              |      |
| PN       | XX  |            |          |          |          |            |          |          |         |              |      |
| XX       | 20-JUN-2002.  |            |          |          |          |            |          |          |         |              |      |
| PD       | XX  |            |          |          |          |            |          |          |         |              |      |
| XX       | 30-NOV-2001; 2001WO-US047576.                                       |            |          |          |          |            |          |          |         |              |      |
| XX       | XX  |            |          |          |          |            |          |          |         |              |      |
| PR       | 12-DEC-2000; 2000US-00735705.                                       |            |          |          |          |            |          |          |         |              |      |
| PR       | 07-MAY-2001; 2001US-00850716.                                       |            |          |          |          |            |          |          |         |              |      |
| PR       | 28-JUN-2001; 2001US-00897778.                                       |            |          |          |          |            |          |          |         |              |      |
| XX       | XX  |            |          |          |          |            |          |          |         |              |      |
| PA       | (CORI-) CORIYA CORP.  |            |          |          |          |            |          |          |         |              |      |

XX Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;  
PI McNeill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedwick TS;  
PI Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;  
XX WFI: 2002-583445/62.  
DR N-PSDB: ABQ92485.  
XX  
XX Novel lung carcinoma polynucleotide sequences and polypeptides encoded by  
PT the polynucleotides, useful in pharmaceutical compositions such as  
PT vaccines and as markers to indicate the presence of lung cancer.  
XX  
XX Claim 9: Page 375-377; 381pp; English.  
XX  
XX The present invention describes isolated human lung carcinoma  
CC polynucleotides (I) and polypeptides (II). (I) and (II) have cytostatic  
CC activity, and can be used in gene therapy and in vaccines. Compositions  
CC comprising (I) or (II) can be used for stimulating an immune response in  
CC a patient and for treating lung cancer in a patient. Oligonucleotides of  
CC (I) can be used for detecting the presence of a cancer in a patient, by  
CC obtaining a biological sample from the patient, contacting the biological  
CC sample with the oligonucleotide, detecting in the sample, an amount of  
CC polynucleotide that hybridises to the oligonucleotide and comparing the

Db 390 ProSerAlaMetThr-----ProProTyrProGlnPheGluGlnSer---GluThrGlu 406  
QY 848 ATGGTGCAGGTGTTATCCCGCCAGCAGCGAGTGGCGGCATCATCGGCAAGAGGGCAG 907  
Db 407 ThrValHisLeuPheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGln 426  
QY 908 CACATCAAAACAGCTCTCCGGTTCAGCGCCCTCCATCAAGATTGCACACCCGAAACA 967  
Db 427 HisLeuLysGlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAla 446  
QY 968 CTTGACTCCAAAGTTCGTATGTTATCATCTGAGCGCGCAGAGGCCCAATTCAGGCT 1027  
Db 447 ProAspAlaLysValArgMetValIleIleThrGlyProGluAlaGlnPheLysAla 466  
QY 1028 CAGGAGAAATCTATGGCAACTCAAGGAGGAGAACTTCTTGTCCTCCAGGAGGAGTC 1087  
Db 467 GlnGlyArgIleTyrGlyLysIleLysGluLysPheValSerProLysGluGluVal 486  
QY 1088 AAGTGGAGACCCATACAGTGTCCAGCATCAGCAGCTGCCGGGTCAITGGCAAGGT 1147  
Db 487 LysLeuGluAlaHisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGly 506  
QY 1148 GGAAGGAGCGTGACGAGTTCAGAAATTCAGCGCAGCTGAGTGGTAGTACCAAGAGAC 1207  
Db 507 GlyLysThrValAsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAsp 526  
QY 1208 CAGACCCCTCATGAGAACAGCAGGTCTATCGTGAATCATCGGACATTTCTATGCCAGT 1267  
Db 527 GlnThrProAspGluAsnAspGlnValValLysIleThrGlyHisPheTyrAlaCys 546  
QY 1268 CAGATGGCTCAACGAGATCCGAGACATCTGCGCCAGGTTAAGCAG---CAGCATCAG 1324  
Db 547 GlnValAlaGlnArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGln 566  
QY 1325 AAGGGA---CAGAGTAAACGAGGCCAGCGCAGGAGGAAG 1360  
Db 567 LysAlaLeuGlnSerGlyProGlnSerArgGly 579  
RESULT 8  
ABP61973  
ID ABP61973 standard; protein; 579 AA.  
AC ABP61973;  
DT 07-OCT-2002 (first entry)  
XX Human lung cancer associated protein sequence SEQ ID NO:446.  
DE Human; lung cancer; lung tumour; cytostatic; gene therapy; vaccine.  
KW Homo sapiens.  
OS  
PN WO200247534-A2.  
PD 20-JUN-2002.  
PP 30-NOV-2001; 2001WO-US047576.  
XX 12-DEC-2000; 2000US-00735705.  
PR 07-MAY-2001; 2001US-00850716.  
PR 28-JUN-2001; 2001US-00897778.  
XX  
XX (CORI-) CORIXA CORP.  
XX Wang T, Wang A, Skeiky VAW, Li SX, Kalos MD, Henderson RA;  
PI McNeill PD, Fanger N, Ratter MW, Durham M, Fanger GR, Vedvick TS;  
PI Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;  
XX  
DR WPI; 2002-583465/62.  
DR N-PSDB; ABQ92483.  
XX  
PI Novel lung carcinoma polynucleotide sequences and polypeptides encoded by

PT the polynucleotides, useful in pharmaceutical compositions such as  
XX vaccines and as markers to indicate the presence of lung cancer.  
PS Claim 9; Page 372-374; 381pp; English.  
XX The present invention describes isolated human lung carcinoma  
CC polynucleotides (I) and polypeptides (II). (I) and (II) have cytostatic  
CC activity, and can be used in gene therapy and in vaccines. Compositions  
CC comprising (I) or (II) can be used for stimulating an immune response in  
CC a patient and for treating lung cancer in a patient. Oligonucleotides of  
CC (I) can be used for detecting the presence of a cancer in a patient, by  
CC obtaining a biological sample from the patient, contacting the biological  
CC sample with the oligonucleotide, detecting in the sample, an amount of  
CC polynucleotide that hybridises to the oligonucleotide and comparing the  
CC amount of polynucleotide that hybridises to the oligonucleotide to a  
CC predetermined cut-off value, and determining the presence of a cancer in  
CC the patient. (I) and (II) are useful in pharmaceutical compositions, e.g.  
CC vaccines. (I) is useful as a marker to indicate the presence or absence  
CC of a cancer such as lung cancer. ABQ92145 to ABQ92486 and ABP61966 to  
CC ABP61992 represent sequences used in the exemplification of the present  
CC invention  
XX  
SQ Sequence 579 AA;

Alignment Scores:  
Pred. No.: 2,36e-144 Length: 579  
Score: 1639.00 Matches: 336  
Percent Similarity: 83.8% Conservative: 44  
Best Local Similarity: 74.1% Mismatches: 57  
Query Match: 52.7% Indels: 16  
DB: Gaps: 9

US-09-270-437D-5 (1-1708) x ABP61973 (1-579)

QY 32 CGGGGAGCCATCATGAAGCTGAATGGCCACAGTTGGAGAACCATGCCCTGAAGGTCTCC 91  
Db 133 ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla 152  
QY 92 TATATCCCGATGACGACATACCA---CAGGGACCTGAGATGGCGCGCAGGG----- 142  
Db 153 TyrIleProAspGluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArg 172  
QY 143 GGCTTGGCTCTCGGGGTGAGCCCGCCAGGGCTCAGCTGGTGGCAGCGGGCCCCAGCC 202  
Db 173 GlyLeuGlyGlnArgGlySerSerArgGlnGlySerPro-----GlySerValSer 189  
QY 203 AAGCAGCAGCAAGTGGACATCCCTCTGGCTCTCTGGTGGTGGTGGTGGTGGTGGTGGT 262  
Db 190 LysGlnLysProCysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAla 209  
QY 263 ATTATGGCAGAGGGGGCCACCATCGCAACATCACAAACACAGCCAGTCCAGATA 322  
Db 210 IleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIle 229  
QY 323 GACGTGCATAGGAGAGAACGAGTGCAGTGCAGTGAAGAACCATCAGTGTGCATCCACC 382  
Db 230 AspValHisArgLysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThr 249  
QY 383 CTTGAGGGTCTCTCTCGCTTGTAGATGATCTTTGGAGATTATCATTAAGAGAGCTAAG 442  
Db 250 ProGluGlyThrSerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGln 269  
QY 443 GACACCAAAACGGCTGACGAGTTCCTGAGATCTCTGAGATCTCTGAGATCTCTGAGG 502  
Db 270 AspIleLysPheThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGly 289  
QY 503 CGTCTCATTTGGCAAGAGGAGCGAACCTGAGAGAGAGTAGAGCAAGATACCGAGCAAAA 562  
Db 290 ArgLeuIleGlyLysGluGlyArgAsnLeuLysLysIleGluGlnAspThrAspThrLys 309  
QY 563 ATCACCATCTCTCTGTCAGACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCT 622  
Db 310 IleThrIleSerProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrVal 329









Pred. No.: 2,36e-144 Length: 579  
Score: 1639.00 Matches: 336  
Percent Similarity: 83.89% Conservative: 44  
Best Local Similarity: 74.17% Mismatches: 57  
Query Match: 52.70% Indels: 16  
DB: 7 Gaps: 9

US-09-270-437D-5 (1-1708) x ADA28539 (1-579)

QY 32 CGGGAGCATCATGAAGCTGAATGGCCACAGTTGGAGAACATGCGCTGAAGTCTCC 91  
DB 133 ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla 152  
QY 92 TACATCCCGGATGAGCAGATAGCA---CAGGGACCTGAGAAATGGCGCCGAGGG----- 142  
DB 153 TyrIleProAspGluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArg 172  
QY 143 GSCITTTGGCTCTCGGGTTCACCCCGCCAGGGCTCACCTGTGGCAGCGGGGCCCGCAGCC 202  
DB 173 GlyLeuGlyGlnArgGlySerArgGlnGlySerPro-----GlySerValSer 189  
QY 203 AAGCAGCAGCAAGTGGACATCCCTTCGGCTCCTGGTCCGCCACCCAGTATGTGGGTGCC 262  
DB 190 LysGlnLysProCysAspLeuProLeuArgLeuValProThrGlnPheValGlyAla 209  
QY 263 ATTATTGGCAGGAGGGGGCCACCATCGCNAACATCACAAACAGACCCAGTCCAGATA 322  
DB 210 IleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIle 229  
QY 323 GACGTGCATAGCAGGAGAACCCAGGTGCAGCTGAAAGGACCATGCTGCATCCACC 382  
DB 230 AspValHisArgLysGluAsnAlaGlyAlaGluLysSerIleThrIleLeuSerThr 249  
QY 383 CTTGAGGCTCTCTCGCTTGAAGTATCTTGAAGTATCTTGAAGTATCTTGAAGTATCTTGAAG 442  
DB 250 ProGluGlyThrSerAlaAlaCysLysSerIleLeuGluIleMethHisLysGluAlaGln 269  
QY 443 GACACCAAAACGGCTGACGAGTTCCTGAAATCTGCGCCCATATAATCTTTGTAGGG 502  
DB 270 AspileLysPheThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGly 289  
QY 503 CGTCTCATTTGCAAGGAGCGAAGCTGAAAGAGGTAGCAAGATACCGAGACAAA 562  
DB 290 ArgLeuIleGlyLysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLys 309  
QY 563 ATCACCATCTCTGTTGCAAGCTTACCTTTACACCTTGAGAGGACCATCTGTG 622  
DB 310 IleThrIleSerProLeuGlnGluLeuThrLeuThrAsnProGluArgThrIleThrVal 329  
QY 623 AAGGGGCGCATCGAGAAATTTGTCAGGCGCCGAGCAGGAAATAATGAGAAAGTTCGGGAG 682  
DB 330 LysGlyAsnValGluThrCysAlaLysAlaGluGluGluIleMethLysLysIleArgGlu 349  
QY 683 GCCTATGAGATGATGGCTGCCATGAGC-----TCTCACCTGATCCCTGCGCTGAAC 736  
DB 350 SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn 369  
QY 737 CTGGCTGCTGTAGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCGG-----CCT 787  
DB 370 LeuAsnAlaLeuGlyLeuPheProProThrSerGlyMetProProThrSerGlyPro 389  
QY 788 CCGAGCAGGTTATCGGGGCTGCTCCCTATAGCTCTTTATGAGGCTCCCGAGCAGGAG 847  
DB 390 ProSerAlaMetThr-----ProProTyrProGlnPheGluGlnSer---GluThrGlu 406  
QY 848 ATGTGTGAGGTTTATCCCGCCAGGAGTGGCGCCCATCATCGCAAGAGGGGCGAG 907  
DB 407 ThrValHisLeuPheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGln 426  
QY 908 CACATCAAAACAGCTCTCCCGGTTTGGCAGGCGCTTCATCAAGATTCACACCCGAAACA 967  
DB 427 HisIleLysGlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAla 446

QY 968 CCTGACTCCAAAGTTCTGATGTTATCATCTACTGACCGCCAGAGGCCCAATTCAGGCT 1027  
DB 447 ProAspAlaLysValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAla 466  
QY 1028 CAGGGAAGAATCTATGGCAAACTCAAGCAGGAGACTTCTTTGGTCCCAAGGAGGAGTG 1087  
DB 467 GlnGlyArgIleTyrGlyLysIleLysGluGlnAsnPheValSerProLysGluGluVal 486  
QY 1088 AAGCTGGAGACCCACATACGTGTGCAGCAGCTGCGCGGTCATTGGCAAAAGGT 1147  
DB 487 LysLeuGluAlaHisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGly 506  
QY 1148 GGAATAACGGTGAACGAGTTCAGATTTCAGCGCAGCTGAGGTGGTAGTACCAAGAGAC 1207  
DB 507 GlyTyrThrValAsnGluLeuGlnAsnLeuSerAlaGluValValProArgAsp 526  
QY 1208 CAGACCCCTGATGAGAACGACGAGTTCATCGTGAATAATCATCGCAGCATTTCTATGCCAGT 1267  
DB 527 GlnThrProAspGluAsnAspGlnValValLysIleThrGlyHisPheTyrAlaCys 546  
QY 1268 CAGATGGTCAACGGAAGATCCGAGACATCTCTGCCCGCAGGTAAAGCAG---CAGCATCAG 1324  
DB 547 GlnValAlaGlnArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGln 566  
QY 1325 AAGGGA---CAGGTAACCGAGCGCCGACGCGCAGGAGGAG 1360  
DB 567 LysAlaLeuGlnSerGlyProProGlnSerArgArgLys 579

RESULT 11  
ADA28438  
ID ADA28438 standard; protein; 579 AA.  
XX  
AC ADA28438;  
XX  
DT 20-NOV-2003 (first entry)  
XX  
DE Human lung tumour protein L5238.  
XX  
KW cancer; lung cancer; gene therapy; vaccine; human;  
XX  
KW lung squamous cell carcinoma.  
XX  
OS Homo sapiens.  
XX  
PN US2003064947-A1.  
XX  
PD 03-APR-2003.  
XX  
PF 30-NOV-2001; 2001US-00007700.  
XX  
PR 18-MAR-1998; 98US-00040802.  
PR 27-JUL-1998; 98US-00123912.  
PR 22-DEC-1998; 98US-00221107.  
PR 02-APR-1999; 99US-00285479.  
PR 17-DEC-1999; 99US-00466396.  
PR 30-DEC-1999; 99US-00476496.  
PR 10-JAN-2000; 2000US-00480884.  
PR 22-FEB-2000; 2000US-00510376.  
PR 04-APR-2000; 2000US-00542615.  
PR 28-JUN-2000; 2000US-00606421.  
PR 02-AUG-2000; 2000US-00630940.  
PR 21-AUG-2000; 2000US-00643597.  
PR 15-SEP-2000; 2000US-00662786.  
PR 09-OCT-2000; 2000US-00685696.  
PR 12-DEC-2000; 2000US-00735705.  
PR 07-MAY-2001; 2001US-00850716.  
PR 28-JUN-2001; 2001US-00897778.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
PI Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;  
PI McNeill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedwick TS;  
PI Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;  
XX

DR WPI; 2003-540798/51.  
 DR N-PSDB; ADA28437.  
 XX New isolated polynucleotides and polypeptides useful for diagnosing,  
 PT preventing and/or treating cancer, particularly lung cancer.  
 PT  
 XX  
 PS Example 2; Page 249-251; 296pp; English.  
 XX

CC The invention describes isolated polynucleotides and polypeptides useful  
 CC for diagnosing, preventing and/or treating cancer, particularly lung  
 CC cancer. A new isolated polynucleotide comprises: any of the 22 fully  
 CC defined nucleotide sequences (e.g. 1012, 900 or 2773 bp) given in the  
 CC specification; complements of the nucleotide sequences cited above; at  
 CC least 10 contiguous residues of the nucleotide sequences cited above; a  
 CC sequence that hybridize to any of the nucleotide sequences under highly  
 CC stringent conditions; a sequence that is at least 75 or 90% identical to  
 CC the above nucleotide sequences; or degenerate variants of the above  
 CC nucleotide sequences. The composition and methods are useful in  
 CC diagnosing, preventing and/or treating cancer, particularly lung cancer,  
 CC in gene therapy and in vaccines. This is the amino acid sequence encoded  
 CC by a human lung tumour cDNA isolated from a lung squamous cell carcinoma  
 CC that may be useful in the diagnosis and treatment of lung cancer and  
 CC other disorders.  
 XX

SQ Sequence 579 AA;

Alignment Scores:

|                        |           |               |     |
|------------------------|-----------|---------------|-----|
| Pred. No.:             | 2,36e-144 | Length:       | 579 |
| Score:                 | 1639.00   | Matches:      | 336 |
| Percent Similarity:    | 83.89%    | Conservative: | 44  |
| Best Local Similarity: | 74.17%    | Mismatches:   | 57  |
| Query Match:           | 52.70%    | Indels:       | 16  |
| DB:                    | 7         | Gaps:         | 9   |

US-09-270-437D-5 (1-1708) x ADA28438 (1-579)

|    |     |  |     |
|----|-----|--|-----|
| Qy | 32  | CGGGAGCCATCATGACCTGAATGGCCACAGTGGAGAACCATGCGCTGAGCTGCC       | 91  |
| Db | 133 | ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGlnAsnPheThrLeuLysValAla | 152 |
| Qy | 92  | TACATCCCGATGACGATAGCA---CAGGGACCTCAGATGGCGCCGAGGG---142      |     |
| Db | 153 | TyrIleProAspGluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArg | 172 |
| Qy | 143 | GGCTTTGGCTCTCGGGTFCAGCCCGCCGAGGGCTCACCTGTGGCAGCGGGGCCAGCC    | 202 |
| Db | 173 | GlyLeuGlyGlnA-GlySerSerArgGlnGlySerPro-----GlySerValSer      | 189 |
| Qy | 203 | AAGCAGCAGCAAGTGACATCCCTTCGGCTCGCTGTCGCCACCCAGTATGTGGTGCC     | 262 |
| Db | 190 | LysGlnLysProCysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAla | 209 |
| Qy | 263 | ATTATTGGCAAGAGGGGGCCACCATCCGCAACATCACAAACAGACCCAGTCCAAGATA   | 322 |
| Db | 210 | IleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIle | 229 |
| Qy | 323 | GACGTGCATAGGAGAGACCGAGGTGCAGCTGAGTAAAAGCCATCATGTGCACTCCACC   | 382 |
| Db | 230 | AspValHisArgLysGluAsnAlaGlyAlaGluLysSerIleThrIleLeuSerThr    | 249 |
| Qy | 383 | CCTCAGGGCTCTCTCCGCTGTGAAGATGATCTTGAGATTATGCAATAAGAGGCTAAG    | 442 |
| Db | 250 | ProGluGlyThrSerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGln | 269 |
| Qy | 443 | GACACCAAAACGGTGCAGAGTTCCCTGAGATCCTGGCCCATATAACTTTGTAGGG      | 502 |
| Db | 270 | AspIleLysPheThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGly | 289 |
| Qy | 503 | CGTCTCATGGCAGGAGGACGCACTCAGAGAGGTAGACAGATACCGACACAAA         | 562 |
| Db | 290 | ArgLeuIleGlyLysGluGlyArgAsnLeuLysLysIleGluGlnAspThrAspThrLys | 309 |
| Qy | 563 | ATCACCATCTCTCGTTGCAAGACCTTACCTTTACAAACCTGAGAGGACCATCACTGTG   | 622 |

|    |      |  |      |
|----|------|--|------|
| Db | 310  | IleThrIleSerProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrVal | 329  |
| Qy | 623  | AAGGGGGCCATCAGAAATGTTGCAGGGCCGAGGAAATAATGAAGAAGTTCCGGAG      | 682  |
| Db | 330  | LysGlyAsnValGluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGlu    | 349  |
| Qy | 683  | GCTATGAGAATGATGGCTGCCATGAGC-----TCTCAGCTGATCCCTGGCTGAC       | 736  |
| Db | 350  | SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn | 369  |
| Qy | 737  | CTGGTCTGTAGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCG-----CCT          | 787  |
| Db | 370  | LeuAsnAlaLeuGlyLeuPheProProThrSerGlyMetProProThrSerGlyPro    | 389  |
| Qy | 788  | CCAGCAGGCTTACTGGGCTGCTCCCTATAGCTCTTATCAGAGTCCCGCAGGAG        | 847  |
| Db | 390  | ProSerAlaMetThr-----ProProTyrProGlnPheGluGlnSer---GluThrGlu  | 406  |
| Qy | 848  | ATGGTGCAGTCTTTATCCCGCCAGGAGTGGCGCCATCATCGGCAAGAGGGCAG        | 907  |
| Db | 407  | ThrValHisLeuPheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGln    | 426  |
| Qy | 908  | CACATCAAAACAGCTCTCCGGTTGCCAGCGCTCCATCAAGATTGCACCCGAAACA      | 967  |
| Db | 427  | HisIleLysGlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAla | 446  |
| Qy | 968  | CCTGACTCCAAAGTTCGTATGTTATCATCTGAGCGCCGAGAGGCCAATTCAGGCT      | 1027 |
| Db | 447  | ProAspAlaLysValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAla | 466  |
| Qy | 1028 | CAGGGAAGAAATCTATGGCAAACTCAAGGAGGAGAACTCTTTTGGTCCCAAGGAGGA    | 1087 |
| Db | 467  | GlnGlyArgIleTyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluVal | 486  |
| Qy | 1088 | AGCTGGAGACCCATACCTGTGCGAGCATCAGCAGTGGCGCGGTCATTTGGCAAGGT     | 1147 |
| Db | 487  | LysLeuGluAlaHisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGly | 506  |
| Qy | 1148 | GGAAAAACCGTCAACGAGTTGCAGAAATTCAGCGAGCTGAGTGGTAGTACCAAGAGAC   | 1207 |
| Db | 507  | GlyLysThrValAsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAsp    | 526  |
| Qy | 1208 | CAGACCCCTGTAGAGAACCGACCGAGTCTATCGTGAATAATCATCGACATTTCTATG    | 1267 |
| Db | 527  | GlnThrProAspGluAsnAspGlnValValLysIleThrGlyHisPheTyrAlaCys    | 546  |
| Qy | 1268 | CAGATGGCTCAACGGAAGATCCGAGACATCCTGGCCCGAGTTAAGCAG---CAGCATCAG | 1324 |
| Db | 547  | GlnValAlaGlnArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGln    | 566  |
| Qy | 1325 | AAGGGA---CAGAGTAACCGAGGCCCGAGCGACGAGGAAG                     | 1360 |
| Db | 567  | LysAlaLeuGlnSerGlyProProGlnSerAspArgLys                      | 579  |

RESULT 12

ABB75048  
 ID ABB75048 standard; protein; 586 AA.  
 XX

AC ABB75048;

XX 01-MAY-2002 (first entry)

XX Human lung tumour L523S recombinant protein sequence SEQ ID NO:427.

XX Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;  
 immune response.

XX Homo sapiens.

XX WO200200174-A2.

XX 03-JAN-2002.

PD

XX 28-JUN-2001; 2001WO-US021065.  
 XX  
 PR 28-JUN-2000; 2000US-00606421.  
 PR 02-AUG-2000; 2000US-006303940.  
 PR 21-AUG-2000; 2000US-00643597.  
 PR 15-SEP-2000; 2000US-00662786.  
 PR 09-OCT-2000; 2000US-00685696.  
 PR 12-DEC-2000; 2000US-00735705.  
 PR 07-MAY-2001; 2001US-00850716.  
 XX  
 PA (CORI-) CORIXA CORP.  
 FI Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;  
 PI McNeill PD, Fanger N, Retter MW, Warnerakis M, Fanger GR;  
 PI Vedvick TS, Carter D, Watanabe Y, Peckham DW;  
 XX  
 DR WPI; 2002-090513/12.  
 DR N-PSDB; ABL49283.  
 XX  
 PT Polynucleotides encoding lung tumor polypeptides, useful for treating  
 PT lung cancer or stimulating an immune response.  
 XX  
 PS Claim 2; Page 354-355; 374pp; English.  
 XX  
 CC The present invention describes human lung tumour proteins. Human lung  
 CC tumour proteins and polynucleotides have cytostatic and immunostimulant  
 CC activities, and can be used in vaccine production. Compositions  
 CC comprising the lung tumour proteins, polynucleotides, antibodies, fusion  
 CC proteins, T cell populations, or antigen presenting cells that express  
 CC the lung tumour proteins are useful for treating lung cancer or  
 CC stimulating an immune response. ABL48959 to ABL49300 and ABL4946 to  
 CC ABL50700 represent sequences used in the exemplification of the present  
 CC invention  
 XX  
 SQ Sequence 586 AA;  
 Alignment Scores:  
 Pred. No.: 2.38e-144 Length: 586  
 Score: 1639.00 Matches: 336  
 Percent Similarity: 83.89% Conservative: 44  
 Best Local Similarity: 74.17% Mismatches: 57  
 Query Match: 52.70% Indels: 16  
 DB: 5 Gaps: 9  
 US-09-270-437D-5 (1-1708) x ABB75048 (1-586)  
 QY 32 CGGGAGCCATCATGAAGCTGAATGGCCACAGTTGGAGAACCATGCCCTGAAGGCTCC 91  
 DB 140 ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla 159  
 QY 92 TACATCCCGATGACAGATAGCA--CAGGACCTGAGATGGCGCCGAGG-----142  
 DB 160 TyrIleProAspGluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArg 179  
 QY 143 GGCCTTGGCTCTCGGGGTGACGCCCGCCGAGGCTCACTGTGGCAGCGGGGCCGAGCC 202  
 DB 180 GlyLeuGlyGlnArgGlySerSerArgGlnGlySerPro-----GlySerValSer 196  
 QY 203 AAGCAGACGACAGTGGACATCCCGCTTGGCTCTCGGTGGTCCGACCCAGATGTGGTGCC 262  
 DB 197 LysGlnLysProCysAspLeuProLeuArgLeuValProThrGlnPheValGlyAla 216  
 QY 263 ATTATTGGCAGGAGGGGGCCACCATCCGCAACATCACAAACAGACCCAGTCCCAAGATA 322  
 DB 217 IleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIle 236  
 QY 323 GACGTGATGAAGAGAGACGAGGTGACGCTGAAAGCCATCAGTGTGACCTCCACC 382  
 DB 237 AspValHisArgLysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThr 256  
 QY 383 CCTGAGGCTGCTCCTCCGCTGTGAAGATATCTTGGAGATTATGATAAGAGGCTAAG 442

DB 257 ProGluGlyThrSerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGln 276  
 QY 443 GACACAAAACGGGTGACGAGTCCCTGAGATCCCTGGCCCATATAACTTTGTAGGG 502  
 DB 277 AspIleLysPheThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGly 296  
 QY 503 CQTCTCATTGGCAAGAGGACGGAACCTGAAGAGGTAGAGCAAGATACCGACACAAA 562  
 DB 297 ArgLeuIleGlyLysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLys 316  
 QY 563 ATCACCATCTCTCTGTCGAAGACCTTACCTTTACAAACCTGAGAGACCATCACTGTG 622  
 DB 317 IleThrIleSerProLeuGlnGluLeuThrLeuTyrAsnProGluAlaGThrIleThrVal 336  
 QY 623 AAGGGGGCCATCGAAGATTGTTGAGGGCCGAGGAGGAATAATGAAGAAGTTCTGGGAG 682  
 DB 337 LysGlyAsnValGluThrCysAlaLysAlaGluGluIleMetLysIleArgGlu 356  
 QY 683 GCCTATGAGATGATGTGGTGGCCATGAGC-----TCTACCTGATCCCTGGCTGAAC 736  
 DB 357 SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn 376  
 QY 737 CTGGCTGCTGAGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCCG-----CCT 787  
 DB 377 LeuAsnAlaLeuGlyLeuPheProThrSerGlyMetProProThrSerGlyPro 396  
 QY 788 CCCAGCAGCTTACTGGGGCTGCTCCCTATAGCTCTTTATGAGGCTCCCGCAGGAG 847  
 DB 397 ProSerAlaMetThr-----ProProThrProGlnPheGluGlnSer---GluThrGlu 413  
 QY 848 ATGGTGCAGGTGTTATCCCGCCGAGCAGTGGCGCCATCATCGCAGAGAGGGCAG 907  
 DB 414 ThrValHisLeuPheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGln 433  
 QY 908 CACATCAAAACAGCTCTCCCGGTTTCCAGCGCTTCATCAAGATGTCACACCCGAAACA 967  
 DB 434 HisIleLysGlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAla 453  
 QY 968 CTGACTCCAAAGTTCTGTTATCATCTGAGCGCCGAGAGGCCCATTCAGGCT 1027  
 DB 454 ProAspAlaLysValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAla 473  
 QY 1028 CAGGGAAGAATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGCTCCCAAGGAGGAAGT 1087  
 DB 474 GlnGlyArgIleTyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluVal 493  
 QY 1088 AAGCTGGAGACCCATACGTGTGCCAGCATCAGCAGTGGCCGGGTCAATTGGCAAGGT 1147  
 DB 494 LysLeuGluAlaHisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGly 513  
 QY 1148 GGAAGAACGGTGAACGAGTTCGAGAATTGAGCGCAGCTGAGGTGGTAGTACCAAGAGAC 1207  
 DB 514 GlyLysThrValAsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAsp 533  
 QY 1208 CAGACCCCTGATGAGACGACACAGGTCTCATCGTGAATAATCATCGGACATTTCTATGCCAGT 1257  
 DB 534 GlnThrProAspGluAsnAspGlnValValLysIleThrGlyHisPheTyrAlaCys 553  
 QY 1258 CAGATGGCTCAACGAGATCCGAGACATCTGTGCCCGCAGGTAAAGAG---CAGCATCAG 1324  
 DB 554 GlnValAlaGlnArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGln 573  
 QY 1325 AAGGGA---CAGAGTAAACCGAGCCGAGCAGGAGGAAG 1360  
 DB 574 LysAlaLeuGlnSerGlyProGlnSerArgArgLys 586  
 RESULT 13  
 ID ABB61968 standard; protein; 586 AA.  
 XX ABB61968  
 AC ABB61968;  
 XX DT 07-OCT-2002 (first entry)

Human lung cancer associated protein sequence SEQ ID NO:427.

Human; lung cancer; lung tumour; cytostatic; gene therapy; vaccine.

Homo sapiens.

WO200247534-A2.

20-JUN-2002.

30-NOV-2001; 2001WO-US047576.

12-DEC-2000; 2000US-00735705.

07-MAY-2001; 2001US-00850716.

28-JUN-2001; 2001US-00897778.

(CORI-) CORIXA CORP.

Wang T, Wang A, Skeiky YAW, Li SX, Kalos WD, Henderson RA;  
Mcneill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedvick TS;  
Carter D, Watanabe Y, Peckham DW, Cai F, Foy TW;

WPI; 2002-583465/62.

N-PSDB; ABQ92469.

Novel lung carcinoma polynucleotide sequences and polypeptides encoded by  
the polynucleotides, useful in pharmaceutical compositions such as  
vaccines and as markers to indicate the presence of lung cancer.

Claim 9; Page 361-362; 381pp; English.

The present invention describes isolated human lung carcinoma  
polynucleotides (I) and polypeptides (II). (I) and (II) have cytostatic  
activity, and can be used in gene therapy and in vaccines. Compositions  
comprising (I) or (II) can be used for stimulating an immune response in  
a patient and for treating lung cancer in a patient. Oligonucleotides of  
(I) can be used for detecting the presence of a cancer in a patient, by  
obtaining a biological sample from the patient, contacting the biological  
sample with the oligonucleotide, detecting in the sample, an amount of  
polynucleotide that hybridises to the oligonucleotide and comparing the  
amount of polynucleotide that hybridises to the oligonucleotide to a  
predetermined cut-off value, and determining the presence of a cancer in  
the patient. (I) and (II) are useful in pharmaceutical compositions, e.g.  
vaccines. (I) is useful as a marker to indicate the presence or absence  
of a cancer such as lung cancer. ABQ92145 to ABQ92486 and ABP61866 to  
ABP61992 represent sequences used in the exemplification of the present  
invention

Sequence 586 AA;

Alignment Scores:

|                        |           |               |     |
|------------------------|-----------|---------------|-----|
| Pred. No.:             | 2,38e-144 | Length:       | 586 |
| Score:                 | 1639.00   | Matches:      | 336 |
| Percent Similarity:    | 83.89%    | Conservative: | 44  |
| Best Local Similarity: | 74.17%    | Mismatches:   | 57  |
| Query Match:           | 52.70%    | Indels:       | 16  |
| DB:                    | 5         | Gaps:         | 9   |

US-09-270-437D-5 (1-1708) x ABP61968 (1-586)

Qy 32 CGGGGAGCCATCATGAGCTGAATCGGCACACCACTGGAGACCATGCTCGAGGCTCC 91

Db 140 ArgGlnAlaLeuAspGlyLeuAsnGlyPheGlnGluGluAsnPheThrLeuGlyValAla 159

Qy 92 TACATCCCCGATGAGCAGATAGCA---CAGGGACCTCGAGATGGCGCCGAGGG----- 142

Db 160 TyrIleProAspGluThrAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArg 179

Qy 143 GGCCTTGGCTCTCGGGGTCAGCCCGCAGAGGCTCACTGTGGCAGCGGGGGCCCGAGCC 202

Db 180 GlyLeuGlyGlnArgGlySerSerArgGlnGlySerPro-----GlySerValSer 196

```
Db 554 GlnValAlaGlnArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGln 573
QY 1325 AAGGGA---CAGAGTAACCCAGGCCAGGACGAGGAG 1360
XX ADA28517
XX ADA28517
AC ADA28517;
XX 20-NOV-2003 (first entry)
XX Recombinant human lung tumour protein L523S.
XX cancer; lung cancer; gene therapy; vaccine; human;
XX lung squamous cell carcinoma.
XX Homo sapiens.
XX US2003064947-A1.
XX 03-APR-2003.
XX 30-NOV-2001; 2001US-00007700.
XX 18-MAR-1998; 98US-00040802.
XX 27-JUL-1998; 98US-00123912.
XX 22-DEC-1998; 98US-00221107.
XX 02-APR-1999; 99US-00285479.
XX 17-DEC-1999; 99US-00466396.
XX 30-DEC-1999; 99US-00476496.
XX 10-JAN-2000; 2000US-00480884.
XX 22-FEB-2000; 2000US-00510376.
XX 04-APR-2000; 2000US-00542615.
XX 28-JUN-2000; 2000US-00606421.
XX 02-AUG-2000; 2000US-00630940.
XX 15-SEP-2000; 2000US-00662786.
XX 09-OCT-2000; 2000US-00685696.
XX 12-DEC-2000; 2000US-00735705.
XX 07-MAY-2001; 2001US-00850716.
XX 28-JUN-2001; 2001US-00897778.
XX (CORI-) CORIXA CORP.
XX Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
XX McNeill PD, Fanger N, Retter WM, Durham M, Fanger GR, Vedvick TS;
XX Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;
XX WPI; 2003-540798/51.
XX New isolated polynucleotides and polypeptides useful for diagnosing,
XX preventing and/or treating cancer, particularly lung cancer.
XX Claim 9; Page 273-275; 296pp; English.
XX The invention describes isolated polynucleotides and polypeptides useful
XX for diagnosing, preventing and/or treating cancer, particularly lung
XX cancer. A new isolated polynucleotide comprises: any of the 22 fully
XX defined nucleotide sequences (e.g. 1012, 900 or 2773 bp) given in the
XX specification; complements of the nucleotide sequences cited above; at
XX least 10 contiguous residues of the nucleotide sequences cited above; a
XX sequence that hybridise to any of the nucleotide sequences under highly
XX stringent conditions; a sequence that is at least 75 or 90% identical to
XX the above nucleotide sequences; or degenerate variants of the above
XX nucleotide sequences. The composition and methods are useful in
XX diagnosing, preventing and/or treating cancer, particularly lung cancer,
XX in gene therapy and in vaccines. This is the amino acid sequence of a
XX recombinant human lung tumour associated protein.
```

```
SQ Sequence 586 AA;
Alignment Scores:
Pred. No.: 2,38e-144 Length: 586
Score: 1639.00 Matches: 336
Percent Similarity: 83.89% Conservatives: 44
Best Local Similarity: 74.17% Mismatches: 57
Query Match: 52.70% Indels: 16
DB: 7 Gaps: 9
US-09-270-437D-5 (1-1708) x ADA28517 (1-586)
QY 32 CGGGAGCCATCATGAAGCTGAATGGCCACAGTTGGAGAACCATGCCCTGAAGGCTCC 91
Db 140 ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla 159
QY 92 TACATCCCCGATGAGCAGATAGCA---CAGGACCTGAGAATGGCGCCGAGGG----- 142
Db 160 TyrIleProAspGluThrAlaAlaGlnAsnProLeuGlnGlnProArgGlyArgArg 179
QY 143 GCGTTTGCTCTCGGGGTACGCCGCCCGCCAGGCTCACCTGTGGCAGCGGGGCCCGACCC 202
Db 180 GlyLeuGlyGlnArgGlySerSerArgGlnGlySerPro-----GlySerValSer 196
QY 203 AAGCAGCAGCAAGTGGACATCCCGCTCGGCTCGCTGGTGGCCACCCAGTATGTGGTGCC 262
Db 197 LysGlnLysProCysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAla 216
QY 263 ATTATTGGCAAGAGGGGGCCACCATCGCAACATCAAAACAGACAGCCAGTCCCAAGATA 322
Db 217 IleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIle 236
QY 323 GACGTGCATAGAGAGAGACGCGAGGTCCAGCTGAAAAGCCATCAGTGTGCATCCACC 382
Db 237 AspValHisArgLysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThr 256
QY 383 CTGAGGGCTGCTCTCCCTGTTGAAGATGATCTTTGGAGATTATGTCATAAGAGGCTAAG 442
Db 257 ProGluGlyThrSerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGln 276
QY 443 GACCCAAACCGGTGACGAGGTTCCTCTGAAGATCTCTGGCCCATATACTTTGTAGGG 502
Db 277 AspIleLysPheThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGly 296
QY 503 CGTCTCATTGGCAAGAGGACGGAACCTGAAGAGGTAGAGCAAGATACCGACACAAAA 562
Db 297 ArgLeuIleGlyLysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLys 316
QY 563 ATCCATCTCTCTGTTGCAAGACCTTACCTTTACAAACCTGAGAGGACCATCCTGTG 622
Db 317 IleThrIleSerProLeuGlnGluLeuThrLeuThrAsnProGluArgThrIleThrVal 336
QY 623 AAGGGGGCCATCGAGAATTGTTGAGGGCCGAGGAGAAATATGAAGAAGTTCCGGAG 682
Db 337 LysGlyAsnValGluThrCysAlaLysAlaGluGluGluIleMetLysLysIleArgGlu 356
QY 683 GCCTATGAGAATGATGTGGCTGCCATGAGC-----TCTCACCTGATCTCCCTGGCTGAAC 736
Db 357 SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn 376
QY 737 CTGGCTGCTGATAGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCG-----CCT 787
Db 377 LeuAsnAlaLeuGlyLeuPheProProThrSerGlyMetProProThrSerGlyPro 396
QY 788 CCCAGCAGCTTACTTGGGGCTGCTCCCTATAGCTCTCTTTATGACAGGTCCTCCGACAGAG 847
Db 397 ProSerAlaMetThr-----ProProTyrProGlnPheGluGlnSer---GluThrGlu 413
QY 848 ATGTGACAGGTGTTTATCCCGCCAGGACGAGTGGGGCCCATCATCGGCAGAGAGGCGAG 907
Db 414 ThrValHisLeuPheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGln 433
QY 908 CACATCAAAACAGCTCTCCCGGTTTTCGACGCGCTCCCATCAAGATTGCACACCCGAAACA 967
```

Db 434 HisLeuGlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAla 453  
 QY 968 CQTGACTCCAAAGTTCGTATGTTATCATCTGACCGCCAGAGGCCCAATTCAGGCT 1027  
 Db 454 ProAspAlaLysValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAla 473  
 QY 1028 CAGGAAGAATCTATGGCAACTCAAGGAGGAGAACTTCTTTGTCCTCCAGGAGGAAGTG 1087  
 Db 474 GlnGlyArgIleTyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluVal 493  
 QY 1088 AAGTGGAGACCCATACAGTGTGCGCAGCATCAGCAGTGTGCGCGGTCTATGGCAAGGT 1147  
 Db 494 LysLeuGluAlaHisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGly 513  
 QY 1148 CGAAACCGGTGAACGAGTTCAGAAATTCAGCGCAGCTGAGGTGGTAGTACCAAGAGAC 1207  
 Db 514 GlyLysThrValAsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAsp 533  
 QY 1208 CAGACCCCTGATGAGACGACGAGTCTATCGTGAATCATCGACATTTCTATGCCAGT 1267  
 Db 534 GlnThrProAspGluAsnAspGlnValValValLysIleThrGlyHisPheTyrAlaCys 553  
 QY 1268 CAGATGGTCAACGGAAGATCCGAGACATCTGCGCCAGGTTAAGCAG---CAGCATCAG 1324  
 Db 554 GlnValAlaGlnArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGln 573  
 QY 1325 AAGGGA---CAGAGTAACGAGCCCGCAGCAGGAGGAAG 1360  
 Db 574 LysAlaLeuGlnSerGlyProProGlnSerArgArgLys 586  
 RESULT 15  
 AAB11328  
 ID AAB11328 standard; protein; 579 AA.  
 XX  
 AC AAB11328;  
 XX  
 DT 21-FEB-2001 (first entry)  
 XX  
 DE Human lung cancer-associated protein L523S.  
 XX  
 KW Lung cancer; therapy; treatment; human; tumor; immunogenic; cytostatic;  
 KW vaccine; detection.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200061612-A2.  
 XX  
 PD 19-OCT-2000.  
 XX  
 PF 03-APR-2000; 2000WO-US008896.  
 XX  
 PR 02-APR-1999; 99US-00285479.  
 PR 17-DEC-1999; 99US-00466396.  
 PR 30-DEC-1999; 99US-00476496.  
 PR 10-JAN-2000; 2000US-00480884.  
 PR 22-FEB-2000; 2000US-00510376.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Wang T, Pan L;  
 XX  
 DR WPI; 2000-628399/60.  
 DR N-PSDB; AAC65900.  
 XX  
 PT Isolated polypeptide comprising an immunogenic portion of a lung tumor  
 PT protein is used for detecting and monitoring progression of lung cancer  
 PT in a patient.  
 XX  
 PS Claim 3; Page 186-188; 261pp; English.  
 XX  
 CC This invention describes a novel isolated polypeptide (I) which  
 CC comprising an immunogenic portion of a lung tumor protein or variant (P2)

CC which have cytostatic activity. The polypeptides and polynucleotides are  
 CC used in compositions and vaccines to inhibit the development of cancer,  
 CC especially lung cancer, in a patient. Methods described in the invention  
 CC can be used to monitor the progression of a cancer by carrying out the  
 CC detection at subsequent time points and comparing the results from the  
 CC different time points. CD4+ and/or CD8+ T-Cells isolated from a patient  
 CC are treated with P2, polynucleotides encoding P2 or antigen presenting  
 CC calls expressing P2 and then administered to the patient to inhibit  
 CC development of cancer  
 XX  
 SQ Sequence 579 AA;

Alignment Scores:  
 Pred. No.: 3,64e-144 Length: 579  
 Score: 1637.00 Matches: 336  
 Percent Similarity: 83.66% Conservative: 43  
 Best Local Similarity: 74.17% Mismatches: 58  
 Query Match: 52.64% Indels: 16  
 DB: 3 Gaps: 9

US-09-270-437D-5 (1-1708) x AAB11328 (1-579)

QY 32 CGGGAGCCATCATGAAGCTGAATGGCCACCATGGAGAACCATGCCCTGAAGGTCTCC 91  
 Db 133 ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla 152  
 QY 92 TACATCCCGATCAGCAGATAGCA---CAGGACCTGAGATGGCGCCGAGG-----142  
 Db 153 TyrIleProAspGluMetAlaAlaGlnGlnAsnP-oLeuGlnGlnProArgGlyArgArg 172  
 QY 143 GGCTTTGGTCTCGGGTCCAGCCCGCAGGCTCACCTGTGCGAGCGGGGCCCCAGCC 202  
 Db 173 GlyLeuGlyGlnArgGlySerSerArgGlnGlySerPro-----GlySerValSer 189  
 QY 203 AAGCAGCAGCAAGTGCACATCCCTTCGGCTCCTGGTCCCGCCACCATGATATGGGTGCC 262  
 Db 190 LysGlnLysProCysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAla 209  
 QY 263 ATTATGGCAAGAGGGGGCCACCATCCGACATCACAAACAGACCCAGTCCAGATA 322  
 Db 210 IleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIle 229  
 QY 323 GAGCTGCATAGGAAGAGAACCGCAGTGCAGTGCAGTGAAGAACCATCAGTGTGCACCTCCAC 382  
 Db 230 AspValHisArgLysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThr 249  
 QY 383 CTGAGGGCTGCTCTCCCTTGTAGATGATCTTGGAGATATGATATGATAAGAGGCTAAG 442  
 Db 250 ProGluGlyThrSerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGln 269  
 QY 443 GACACCAAAACGGCTGACGAGGTCCCTGAAGATCCTGGCCCAATAAATCTTGTAGGG 502  
 Db 270 AspIleLysPheThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGly 289  
 QY 503 CGTCTATTGGCAAGGAAGCGAACCTGAAGAGGTAGAGCAAGATACCGAGACAAA 562  
 Db 290 ArgLeuIleGlyLysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLys 309  
 QY 563 ATCACCATCTCTCTGTCAGACCTTACCTTTACACCTTACACCTTGAGAGACCATCATCTG 622  
 Db 310 IleThrIleSerProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrVal 329  
 QY 623 AAGGGGGCCATCGAGAAATTTGTTGACGGCCGACGAGGAAATAATGAAGAAGTTCGGGAG 682  
 Db 330 LysGlyAsnValGluThrCysAlaLysAlaGluGluGluIleMetLysLysIleArgGlu 349  
 QY 683 GCCTATGAGAAATGATGTGCTGCCATGAGC-----TCTACCTGATCCCTGGCTGTAAC 736  
 Db 350 SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn 369  
 QY 737 CTGGGTGCTGTAGGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCCG-----CCT 787  
 Db 370 LeuAsnAlaLeuGlyLeuPheProThrSerGlyMetProProProThrSerGlyPro 389

```
QY 788 CCCAGCAGCGTTACTGGGGTGCTCCCTATAGTCTCCTTTATGAGGCTCCCGAGCAGGAG 847
Db |||:|||||
390 ProSerAlaMetThr-----ProProTyrProGlnPheGluGlnSer--GluThrGlu 406
QY 848 ATGGTGCAGGTGTTTATCCCGCCAGCGCGGCGCCATCTCCGCAAGAGAGGGGCGAG 907
Db |||:|||||
407 ThrValHisGlnPheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGln 426
QY 908 CACATCAAAACAGCTCTCCCGGTTTGGCCAGGCGCTCCCATCAAGATTGCACCACCCGAAACA 967
Db |||:|||||
427 HisIleLysGlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAla 446
QY 968 CCTGACTCCAAAGTTCGTATGTTATCATCTGACCGCCGAGAGGCCCAATTCAGGCT 1027
Db |||:|||||
447 ProAspAlaLysValArgMetValIleIleThrGlyProGluAlaGlnPheLysAla 466
QY 1028 CAGGGAAGATCTATGGCAACTCAAGGAGAGAACTTCTTGGTCCCAAGGAGGAAGTG 1087
Db |||:|||||
467 GlnGlyArgIleTyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluVal 486
QY 1088 AAGCTGGAGACCCACATACGTGTCCAGCATCAGCAGCTGCGCGGTTCATTGGCAAAGGT 1147
Db |||:|||||
487 LysLeuGluAlaHisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGly 506
QY 1148 GGAAAAACGGTGAACGAGTTGCAGAAATTGACGGCAGCTGAGGTGGTAGTACCAAGAGAC 1207
Db |||:|||||
507 GlyLysThrValAsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgASP 526
QY 1208 CAGACCCCTGATGAGAACGACCCAGGTTCATCGTGAATAATCATCGGACATTTCTATGCCAGT 1267
Db |||:|||||
527 GlnThrProAspGluAsnAspGlnValValLysIleThrGlyHisPheTyrAlaCys 546
QY 1268 CAGATGGCTCAACGGAAGATCCGAGACATCTGCGCCCGAGGTTAAGCAG---CAGCATCAG 1324
Db |||:|||||
547 GlnValAlaGlnArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGln 566
QY 1325 AAGGGA---CAGAGTAACACAGGCCCGCCAGGCACGGAGGAAG 1360
Db |||:|||||
567 LysAlaLeuGlnSerGlyProProGlnSerArgArgLys 579
```

Search completed: July 13, 2004, 12:08:40  
Job time : 108 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: July 13, 2004, 12:07:12 ; Search time 22 Seconds  
(without alignments)  
8016.105 Million cell updates/sec

Title: US-09-270-437D-5  
Perfect score: 3110  
Sequence: 1 aggaagctgcgcacccgc.....attcttcagggtttaaaaa 1708

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 778828

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+n2p.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US09270437/runat\_13072004\_121922\_9641/app\_query.fasta\_1.1863  
-DB=Issued Patents AA -QPMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09270437.scgn\_1\_33@runat\_13072004\_121922\_9641 -NCFU=6 -ICPU=3  
-NO MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A.COMB.pcp.\*  
2: /cgn2\_6/ptodata/2/iaa/5B.COMB.pcp.\*  
3: /cgn2\_6/ptodata/2/iaa/6A.COMB.pcp.\*  
4: /cgn2\_6/ptodata/2/iaa/6B.COMB.pcp.\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS.COMB.pcp.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query | Score | Match | Length | DB                 | ID | Description       |
|------------|-------|-------|-------|--------|--------------------|----|-------------------|
| 1          | 2208  | 71.0  | 577   | 3      | US-09-261-855-2    |    | Sequence 2, Appli |
| 2          | 1639  | 52.7  | 579   | 4      | US-09-643-597-348  |    | Sequence 348, App |
| 3          | 1639  | 52.7  | 579   | 4      | US-09-542-615A-348 |    | Sequence 348, App |
| 4          | 1639  | 52.7  | 579   | 4      | US-09-606-421B-348 |    | Sequence 348, App |
| 5          | 1637  | 52.6  | 579   | 4      | US-09-643-597-176  |    | Sequence 176, App |
| 6          | 1637  | 52.6  | 579   | 4      | US-09-480-884A-176 |    | Sequence 176, App |
| 7          | 1637  | 52.6  | 579   | 4      | US-09-542-615A-176 |    | Sequence 176, App |
| 8          | 1637  | 52.6  | 579   | 4      | US-09-606-421B-176 |    | Sequence 176, App |
| 9          | 248.5 | 8.0   | 643   | 1      | US-08-021-608D-10  |    | Sequence 10, Appl |
| 10         | 248.5 | 8.0   | 643   | 1      | US-08-726-160-10   |    | Sequence 10, Appl |
| 11         | 248.5 | 8.0   | 643   | 5      | PCT-US94-01782-10  |    | Sequence 10, Appl |
| 12         | 245   | 7.9   | 49    | 3      | US-09-261-855-18   |    | Sequence 18, Appl |

|    |       |     |      |   |                      |  |                   |
|----|-------|-----|------|---|----------------------|--|-------------------|
| 13 | 245   | 7.9 | 644  | 1 | US-08-021-608D-2     |  | Sequence 2, Appli |
| 14 | 245   | 7.9 | 644  | 1 | US-08-726-160-2      |  | Sequence 2, Appli |
| 15 | 245   | 7.9 | 644  | 5 | PCT-US94-01782-2     |  | Sequence 2, Appli |
| 16 | 242   | 7.8 | 590  | 1 | US-08-021-608D-8     |  | Sequence 8, Appli |
| 17 | 242   | 7.8 | 590  | 1 | US-08-726-160-8      |  | Sequence 8, Appli |
| 18 | 242   | 7.8 | 590  | 5 | PCT-US94-01782-8     |  | Sequence 8, Appli |
| 19 | 236   | 7.6 | 48   | 3 | US-09-261-855-20     |  | Sequence 20, Appl |
| 20 | 229   | 7.4 | 47   | 3 | US-09-261-855-17     |  | Sequence 17, Appl |
| 21 | 228   | 7.3 | 47   | 3 | US-09-261-855-19     |  | Sequence 19, Appl |
| 22 | 222   | 7.1 | 49   | 3 | US-09-261-855-22     |  | Sequence 22, Appl |
| 23 | 221.5 | 7.1 | 530  | 1 | US-08-187-793-4      |  | Sequence 4, Appli |
| 24 | 219   | 7.0 | 48   | 3 | US-09-261-855-24     |  | Sequence 24, Appl |
| 25 | 214   | 6.9 | 47   | 3 | US-09-261-855-21     |  | Sequence 21, Appl |
| 26 | 191   | 6.7 | 47   | 3 | US-09-261-855-23     |  | Sequence 23, Appl |
| 27 | 176.5 | 6.7 | 1418 | 4 | US-09-252-991A-32367 |  | Sequence 32367, A |
| 28 | 174   | 5.6 | 343  | 1 | US-08-187-793-2      |  | Sequence 2, Appli |
| 29 | 162.5 | 5.2 | 697  | 4 | US-09-252-991A-24009 |  | Sequence 24009, A |
| 30 | 157   | 5.0 | 782  | 4 | US-09-252-991A-30464 |  | Sequence 30464, A |
| 31 | 154.5 | 5.0 | 243  | 1 | US-08-021-608D-6     |  | Sequence 6, Appli |
| 32 | 154.5 | 5.0 | 243  | 1 | US-08-726-160-6      |  | Sequence 6, Appli |
| 33 | 154.5 | 5.0 | 243  | 5 | PCT-US94-01782-6     |  | Sequence 6, Appli |
| 34 | 154   | 5.0 | 577  | 4 | US-09-252-991A-19178 |  | Sequence 19178, A |
| 35 | 151.5 | 4.9 | 488  | 4 | US-09-252-991A-19034 |  | Sequence 19034, A |
| 36 | 151.5 | 4.9 | 489  | 4 | US-09-252-991A-16940 |  | Sequence 16940, A |
| 37 | 151   | 4.9 | 563  | 4 | US-09-252-991A-28458 |  | Sequence 28458, A |
| 38 | 150   | 4.8 | 904  | 4 | US-09-976-594-615    |  | Sequence 615, App |
| 39 | 145   | 4.7 | 957  | 4 | US-09-252-991A-20408 |  | Sequence 20408, A |
| 40 | 144.5 | 4.6 | 2294 | 4 | US-09-252-991A-17231 |  | Sequence 17231, A |
| 41 | 143.5 | 4.6 | 878  | 4 | US-09-252-991A-31974 |  | Sequence 31974, A |
| 42 | 142.5 | 4.6 | 774  | 4 | US-09-252-991A-16789 |  | Sequence 16789, A |
| 43 | 141.5 | 4.5 | 783  | 4 | US-09-252-991A-18035 |  | Sequence 18035, A |
| 44 | 141   | 4.5 | 443  | 2 | US-08-935-450-6      |  | Sequence 6, Appli |
| 45 | 141   | 4.5 | 443  | 4 | US-08-952-899-2      |  | Sequence 2, Appli |

ALIGNMENTS

RESULT 1

US-09-261-855-2  
; Sequence 2, Application US/09261855A  
; Patent No. 6255055  
; GENERAL INFORMATION:  
; APPLICANT: Ross, Jeffrey  
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN  
; FILE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE  
; FILE REFERENCE: 960296.95131  
; CURRENT APPLICATION NUMBER: US/09/261,855A  
; CURRENT FILING DATE: 1999-03-03  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 577  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-261-855-2

Alignment Scores:  
Pred. No.: 8.33e-200 Length: 577  
Score: 2208.00 Matches: 438  
Percent Similarity: 98.65% Conservative: 1  
Best Local Similarity: 98.43% Mismatches: 4  
Query Match: 71.00% Indels: 2  
DB: 3 Gaps: 1

US-09-270-437D-5 (1-1708) x US-09-261-855-2 (1-577)

|    |     |   |     |
|----|-----|---|-----|
| Qy | 32  | CGGGAGCCATCATGAGCTGAATGCCACCATGTTGGAGAACCATGCCCTGAAGTCTCC   | 91  |
| Db | 133 | AGGlnAlaIleMetIysLeuAsnGlyHisGlnLeuGluAsnHisAlaLeuIysValSer | 152 |
| Qy | 92  | TACATCCCGGATGACAGATACACAGGACCTGAGATGGCGCGAGGGGCTTGGC        | 151 |
| Db | 153 | TyrIleProAspGluGlnIleThrGlnGlyProGluAsnGlyArgGlyGlyPheGly   | 172 |

152 QY TCTGGGTCAGCCCGCAGGGCTCACCTGTGTCAGCGGGGGCCCGCAGCAGCAG 211  
Db SerArgGlyGlnProArgGlnGlySerProValAlaAlaGlyAlaProAlaLysGlnGln 192  
QY 212 CAAGTGCACATCCCGCTTCTGGCTCTGGTCCGCCACCCAGTATGTGGTGCATATTGGC 271  
Db ProValAspIleProLeuArgLeuValProThrGlnTrpValGlyAlaIleGly 212  
QY 272 AAGGAGGGGGCCACATCCCGCAATCACAAAACAGACCAGTCCAAAGATAGAGCTGCAT 331  
Db LysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHis 232  
QY 332 AGGAAGGAGAACCGAGTGCAGCTGAAAGCCATCATGTGTGCATCTCCACCCCTGAGGGC 391  
Db ArgLysGluAsnAlaGlyAlaAlaGluLysAlaIleSerValHisSerThrProGluGly 252  
QY 392 TGCTCCTCCGCTTGTAGATGATCTTGAGATATGCTAAGAGGCTAAGGACACAAA 451  
Db CysSerSerAlaCysLysMetIleGluGluIleValHisGluAlaLysAspThrLys 272  
QY 452 ACGCTGACGAGTTCCTGAGATCCTGGCCCATTAATCTTGTAGGCGCTCTCAT 511  
Db ThrAlaAspGluValProLeuLysIleLeuAlaHisAsnAspValGlyArgLeuIle 292  
QY 512 GGCAGGAGGACGACGACCTGAGAGGTAGAGATACCGAGATACCGACAAAATCACCATC 571  
Db GlyLysGluGlyArgAsnLeuLysValGluGlnAspThrGluThrLysIleThrIle 312  
QY 572 TCCTCGTTGCAAGACCTTACCTTTACACCTGAGAGGACCATCACTGTGAAGGGGGCC 631  
Db SerSerLeuGlnAspLeuThrLeuTyAsnProGluArgThrIleThrValLysGlyAla 332  
QY 632 ATCAGAAATCTTTCAGGGCCGACAGGAATATAGAAAGTTCGGGAGGCTATGAG 691  
Db IleGluAsnCysCysArgAlaGluGlnIleMetLysLysValArgGluAlaTyArgLys 352  
QY 692 AATGATGGCTGCATGAGC-----TCTCACCTGATCCTCGCCCTGAACCTGGCTGCT 745  
Db AsnAspValAlaAlaMetSerLeuGlnSerHisLeuIleProGlyLeuAsnLeuAlaAla 372  
QY 746 GTAGCTCTTTCCAGCTTATCCAGCGAGTCCCGCTCCCGCTCCCGACGCTTCTGGG 805  
Db ValGlyLeuPheProAlaSerSerAlaValProProProProSerSerValThrGly 392  
QY 806 GCTGCTCCCTATAGCTCTTATGCTAGCTCCCGACGAGGATGTCAGAGGTTTATC 865  
Db AlaAlaProTySerSerPheMetGluAlaProGluGlnMetValGlnValPheIle 412  
QY 866 CCCGCCAGGAGTGGCGCCATCATCGCAAGAGGGGCGAGCAGATCAACAGCTCTCC 925  
Db ProAlaGlnAlaValGlyAlaIleIleGlyLysLysGlyGlnHisIleLysGlnLeuSer 432  
QY 926 CGGTTCCAGCGCTCCATCAAGATTGCACCCCGAAGACCTGACTCCAAAGTTCGT 985  
Db ArgPheAlaSerAlaSerIleLysIleAlaProProGluThrProAspSerLysValArg 452  
QY 986 ATGTTTATCATCACTGGACCGCAGAGGCGCCCAATTCAGGCTCAGGGAAGAACTTATGGC 1045  
Db MetValValIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyGly 472  
QY 1046 AAATCAAGAGGAGAACTCTTTGGTCCCAAGAGGAAGTGAAGCTGGAGCCACATA 1105  
Db LysLeuLysGluGluAsnPhePheGlyProLysGluGluValLysLeuGluThrHisIle 492  
QY 1106 CGTGTCCAGCATCAGCAGCTGGCGGCTCATTTGGCAAGAGTTCGAAAACGGTGAACGAG 1165  
Db ArgValProAlaSerAlaAlaGlyArgValIleGlyLysGlyGlyLysThrValAsnGlu 512  
QY 1166 TTGCAAGATTGACGGAGCTGAGGTGGTAGTACCAAGAGACCCAGACCCCTGATGAGAC 1225  
Db LeuGlnAsnLeuThrAlaAlaGluValValProArgAspGlnThrProAspGluAsn 532

1226 QY GACCAAGGTCACTCGTCAAAATCATCGACATTTCTATGTCACATCGCTCAACGGAG 1285  
Db AspGlnValIleValLysIleIleGlyHisPheTyAlaSerGlnMetAlaGlnArgLys 552  
QY 1286 ATCCGAGACATCCTGGCCCGAGTTAAGCAGCAGCATCAGAAGGACAGAGTAACAGGCC 1345  
Db IleArgAspIleLeuAlaGlnValLysGlnHisGlnLysGlyGlnSerAsnLeuAla 572  
QY 1346 CAGGCACGAGGAG 1360  
Db 573 GlnAlaArgArgLys 577  
RESULT 2  
US-09-643-597-348  
; Sequence 348, Application US/09643597  
; Patent No. 6426072  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongdong  
; APPLICANT: Pan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: McNeill, Patricia D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C11  
; CURRENT APPLICATION NUMBER: US/09/643,597  
; CURRENT FILING DATE: 2000-08-21  
; NUMBER OF SEQ ID NOS: 369  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 348  
; LENGTH: 579  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-643-597-348

Alignment Scores:  
Pred. No.: 4,72e-146 Length: 579  
Score: 1639.00 Matches: 336  
Percent Similarity: 83.89% Conservative: 44  
Best Local Similarity: 74.17% Mismatches: 57  
Query Match: 52.70% Indels: 16  
DB: 4 Gaps: 9  
US-09-270-437D-5 (1-1708) x US-09-643-597-348 (1-579)  
QY 32 CGGGGAGCATCATCAAGCTGAATGGCCACCCAGTTGGAGAACCATGCGCTCAAGGTCTCC 91  
Db 133 ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla 152  
QY 92 TACATCCCGATGACAGATAGCA---CAGGACCTGAGAAATGGCGCCGAGGG--- 142  
Db 153 TyrlleProAspGluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArg 172  
QY 143 GCCTTTGGCTCTCGGGGTACGCCCGCCAGGCTCACCTGGCGAGCGGGGGCCCGAC 202  
Db 173 GlyLeuGlyGlnArgGlySerArgGlnGlySerPro-----GlySerValSer 189  
QY 203 AAGCAGCAGCAAGTGCACATCCCTTCGGTCTCTGGTCCGCCACCCAGTATGCGGTGCC 262  
Db 190 LysGlnLysProCysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAla 209  
QY 263 ATTATTGCAAGGAGGGGGCCACCATCCGCAACATCACAACAGACCCAGTCCCAAGATA 322  
Db 210 IleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIle 229  
QY 323 GACGTGCATAGAGGAGACCGCAGGTGCAGCTGAAAAAGCCATCAGTGTGCTCCACCC 382

230 AspValHisArgLysGluAsnAlaGlyAlaGluLysSerIleThrIleLeuSerThr 249  
383 CCTGAGGGCTGCTCTCGCTGTTGAAGATGATCTTGAGATATTGCTAAGAAGGCTAAG 442  
250 ProGluGlyThrSerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGln 269  
443 GACACCAAAACGGCTGACGAGGTTCCCTGAAGATCCTGGCCATATAACTTTGAGGG 502  
270 AspileLysPheThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGly 289  
503 CGTCTCATTTGGCAAGAGGACGGAACCTGAAGAGGTAGAGCAAGATACCGACACAAA 562  
290 ArgLeuIleGlyLysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLys 309  
563 ATCACCATCTCTCGTTGCAAGACCTTACCTTTACACCCCTGAGAGGACCATCACTGTG 622  
310 IleThrIleSerProLeuGlnGluLeuThrLeuThrLeuThrLeuThrVal 329  
623 AAGGGGCGCATCGAGAATTGTTGAGGGCCGAGCAGGAATAATGAAGAAGTTCCGGAG 682  
330 LysGlyAsnValGluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGlu 349  
683 GCCTATGAGATGATGGCTGGCTGCCATGAGC-----TCTACCTGATCCCTGGCCTGAAC 736  
350 SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn 369  
737 CTGGCTGCTGTAGGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCG-----CCT 787  
370 LeuAsnAlaLeuGlyLeuPheProProThrSerGlyMetProProThrSerGlyPro 389  
788 CCCAGCAGCGTTACTGGGGTGCTCCCTATAGCTCTTTATGAGCGCTCCCGAGCAGGAG 847  
390 ProSerAlaMetThr-----ProProTyrProGlnPheGluGlnSer---GluThrGlu 406  
848 ATGTGTGAGGTGTTTATCCCGCCAGCGAGTGGCGCCATCATCGGCAAGAGAGGGGCGAG 907  
407 ThrValHisLeuPheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGln 426  
908 CACATCAAAACAGCTCTCCGCTTTGGCCAGCGCTCCATCAAGATGTCACCCAGCAACA 967  
427 HisIleLysGlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAla 446  
968 CTTGACTCCAAAGTTCGTATGGTTATCATCACTGACCGCGCAGAGGCCCAATTCAGGCT 1027  
447 ProAspAlaLysValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAla 466  
1028 CAGGGAAGAATCTATGGCAAACTCAAGGAGAGAACTTCTTTGTCCTCAAGGAGGAAGTG 1087  
467 GlnGlyArgIleTyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluVal 486  
1088 AAGCTGAGACCCCATACAGTGTGCCAGCATCAGCAGCTGCGCGGTTCATTGGCAAGGT 1147  
487 LysLeuGluAlaHisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGly 506  
1148 GGAAAAACGGTGAACGAGTTGCAGAAATTTGACGCGCAGCTGAGTGGTAGTACCAAGAGAC 1207  
507 GlyLysThrValAsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAsp 526  
1208 CAGACCCCTGATGAGAACGACCAAGGTTCATCGTGAATAATCATCGACATTTCTATGCCAGT 1267  
527 GlnThrProAspGluAsnAspGlnValValLysIleThrGlyHisPheTyrAlaCys 546  
1268 CAGATGCTCAACGAGATCCGAGACATCTCTGCCCGCAGGTTAAGCAG-----CAGCATCAG 1324  
547 GlnValAlaGlnArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGln 566  
1325 AAGGGA---CAGAGTAACACGAGCCGAGCGCAGGAGGAAG 1360  
567 LysAlaLeuGlnSerGlyProProGlnSerArgArgLys 579

RESULT 3

US-09-542-615A-348

; Sequence 348, Application US/09542615A

; Patent No. 6518256  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy A.  
; APPLICANT: Fanger, Gary R.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C8  
; CURRENT APPLICATION NUMBER: US/09/542,615A  
; CURRENT FILING DATE: 2000-04-14  
; NUMBER OF SEQ ID NOS: 350  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 348  
; LENGTH: 579  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-542-615A-348  
  
Alignment Scores:  
Pred. No.: 4,72e-146 Length: 579  
Score: 1639.00 Matches: 336  
Percent Similarity: 83.89% Conservative: 44  
Best Local Similarity: 74.17% Mismatches: 57  
Query Match: 52.70% Indels: 16  
DB: Gaps: 9  
  
US-09-270-437D-5 (1-1708) x US-09-542-615A-348 (1-579)  
  
QY 32 CGGGGAGCCATCATGAAGCTGAATGGCCACCACTGGAGAACCATGCCCTGAGGTCTCC 91  
Db 133 ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla 152  
QY 92 TACATCCCGATGAGCAGATAGCA---CAGGACCTGAGATGGCGCGGAGGG----- 142  
Db 153 TyrIleProAspGluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArg 172  
QY 143 GCGTTTGGCTCTCGGGGTCAGCCCGCCAGGGCTCACCTGTGGCAGCGGGGCCCGCAGC 202  
Db 173 GlyLeuGlyGlnArgGlySerSerArgGlnGlySerPro-----GlySerValSer 189  
QY 203 AAGCAGCAGCAAGTGGACATCCCTTCGGCTCTGCTGGTCCGCCACCATGATGTTGGTGC 262  
Db 190 LysGlnLysProCysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAla 209  
QY 263 ATTATTGGCAAGGAGGGGCCACCATCCGACATCAGCAACATCAGAAACAGACCCAGTCA 322  
Db 210 IleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIle 229  
QY 323 GACGTGCATAGGAAGAGAACCGAGTGCAGCTGAGTGAAGAACCCATCAGTGTGCACCTCC 382  
Db 230 AspValHisArgLysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThr 249  
QY 383 CTTGAGGGCTGCTCTCCGCTTGTAAAGATGATCTTGGAGATTATGCATAAAGAGGCTAAG 442  
Db 250 ProGluGlyThrSerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGln 269  
QY 443 GACACCAAAACGGCTGACGAGGTTCCCTGAAAGATCTCGGCCCATATAAATCTTGTAGG 502  
Db 270 AspileLysPheThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGly 289  
QY 503 CTTCTCATTTGGCAAGAGGACGGAACCTGAAGAGGTAGAGCAAGATACCGACACAAA 562  
Db 290 ArgLeuIleGlyLysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLys 309  
QY 563 ATCACCATCTCTCGTTGCAAGACCTTACCTTTTACACCCCTGAGAGGACCATCACTGTG 622  
Db 310 IleThrIleSerProLeuGlnGluLeuThrLeuThrLeuThrLeuThrVal 329  
QY 623 AAGGGGCGCATCGAGAATTGTTGAGGGCCGAGCAGGAATAATGAAGAAGTTCCGGAG 682



QY 908 CACATCAAAAGCTCTCCGGTGTTCAGAGCCCTCCATCAAGATTGCAACACCCGAAACA 967  
 Db 427 HisIleLysGlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAla 446  
 QY 968 CCTCACTCCAAAGTTCGTATGTTATCATCTGACCGCCAGAGGCCCAATTCAGGGCT 1027  
 Db 447 ProAspAlaLysValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAla 466  
 QY 1028 CAGGGAAGAATCTATGGCAAACTCAAGAGGAGAACTTCTTTGGTCCCAAGGAGAGTG 1087  
 Db 467 GlnGlyArgIleTyrGlyLysIleLysGluAsnPheValSerProLysGluGluVal 486  
 QY 1088 AAGCTGGAGACCCACATACGTCGCCAGCATCACAGCTGCCGGTTCATGGCAAGGT 1147  
 Db 487 LysLeuGluAlaHisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGly 506  
 QY 1148 GGAATAACGGTGAACGAGTTGACAAATTCAGCGCAGCTGAGGTGGTAGTACCAAGAGAC 1207  
 Db 507 GlyLysThrValAsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAsp 526  
 QY 1208 CAGACCCCTGATGAGAACGACCATGTCATCGTGGAAATCATCGGACATTTCTATGCCAGT 1267  
 Db 527 GlnThrProAspGluAsnAspGlnValValLysIleThrGlyHisPheTyrAlaCys 546  
 QY 1268 CAGATGGCTCAACGGAAGATCCGAGACATCTCGCCAGGTAAAGCAG---CAGCATCAG 1324  
 Db 547 GlnValAlaGlnArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGln 566  
 QY 1325 AAGGGA---CAGATTAACGAGCCGACGACGAGGAG 1360  
 Db 567 LysAlaLeuGlnSerGlyProProGlnSerArgArgLys 579

## RESULT 5

US-09-643-597-176  
 ; Sequence 176, Application US/09643597  
 ; Patent No. 6426072  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Tongtong  
 ; APPLICANT: Fan, Liqun  
 ; APPLICANT: Kalos, Michael D.  
 ; APPLICANT: Baqur, Chaitanya S.  
 ; APPLICANT: Hosken, Nancy  
 ; APPLICANT: Fanger, Gary R.  
 ; APPLICANT: Li, Samuel X.  
 ; APPLICANT: Wang, Aijun  
 ; APPLICANT: Skeiky, Yasir A.W.  
 ; APPLICANT: Henderson, Robert A.  
 ; APPLICANT: McNeill, Patricia D.  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
 ; FILE REFERENCE: 210121.455C11  
 ; CURRENT APPLICATION NUMBER: US/09/643,597  
 ; CURRENT FILING DATE: 2000-08-21  
 ; NUMBER OF SEQ ID NOS: 369  
 ; SOFTWARE: Fast-Seq for Windows Version 3.0  
 ; SEQ ID NO 176  
 ; LENGTH: 579  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-643-597-176

Alignment Scores:  
 Pred. No.: 7,3e-146 Length: 579  
 Score: 1637.00 Matches: 336  
 Percent Similarity: 83.66% Conservative: 43  
 Best Local Similarity: 74.17% Mismatches: 58  
 Query Match: 52.64% Indels: 16  
 DB: 4 Gaps: 9

US-09-270-437D-5 (1-1708) x US-09-643-597-176 (1-579)

QY 32 CGGGAGCCCATCATGAAGCTGAATGGCCACGAGTGGAGAACCATGCCCTGAAGTCTCC 91

Db 133 ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla 152  
 QY 92 TACATCCCCGATCAGCAGATAGCA---CAGGACCTGAGAAATGGCGCCGAGG----- 142  
 Db 153 TyrIleProAspGluMetAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArg 172  
 QY 143 GGCCTTGGCTCTCGGGTTCAGCCCGCAGGGCTCACCTGTGGCAGCGGGGGCCCCAGCC 202  
 Db 173 GlyLeuGlyGlnArgGlySerSerArgGlnGlySerPro-----GlySerValSer 189  
 QY 203 AAGCAGCAGCAAGTGACATCCCCCTTGGCTCTCTGGTCCGCCACCCAGTATGTGGGTGCC 262  
 Db 190 LysGlnLysProCysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAla 209  
 QY 263 ATATTGGCAAGAGGGGGCCACCATCCGCAACATACAAAAACAGACCCAGTCCAGATA 322  
 Db 210 IleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIle 229  
 QY 323 GACGTGCATAGAGAGAGAACCGCAGCTGCAGCTGAGAAAGCCATCAGTGTGACTCCACC 382  
 Db 230 AspValHisArgLysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThr 249  
 QY 383 CTGTAGGGCTGCTCTCCGCTTGTAAAGATGATCTTTGGAGATTATCATAAAGAGGCTAAG 442  
 Db 250 ProGluGlyThrSerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGln 269  
 QY 443 GACACCAAAACGGCTGACAGGTTCCCTGAGAGATCTCGGCCCATATAAATCTTTAGGG 502  
 Db 270 AspIleLysPheThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGly 289  
 QY 503 CGTCTCATGGCAAGAGGACGAACTGAAAGAGGTAGACAAAGATACCCAGACAAAC 562  
 Db 290 ArgLeuIleGlyLysGluGlyArgAsnLeuLysLysIleGluGlnAspThrAspThrLys 309  
 QY 563 ATCACCATCTCTCGTTGCAAGACCTTACCTCTTCAACCTTGAGAGACCATCATCTGTG 622  
 Db 310 IleThrIleSerProLeuGlnLeuThrLeuTyrAsnProGluArgThrIleThrVal 329  
 QY 623 AAGGGGGCCATCGAGAAATTTGTCAGGGCCGAGAGGAAATAATCAAGAAAGCTCGGAG 682  
 Db 330 LysGlyAsnValGluThrCysAlaLysAlaGluGluGluIleMetLysLysIleArgGlu 349  
 QY 683 GCCTATGAGAATGATGTGCTGCCATGAGC-----TCTCCTCATCTCCCTGGCTGAAC 736  
 Db 350 SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn 369  
 QY 737 CTGGCTGTGTAGTCTTTTCCAGCTTCATCCAGGCGAGTCCCGCG-----CCT 787  
 Db 370 LeuAsnAlaLeuGlyLeuPheProThrSerGlyMetProProProThrSerGlyPro 389  
 QY 788 CCAGCAGCGTTACTCGGGCTGCTCCTATAGCTCTTTATGACAGCTCCCGCAGCAGGAG 847  
 Db 390 ProSerAlaMetThr-----ProProTyrProGlnPheGluGlnSer---GluThrGlu 406  
 QY 848 ATGGTGCAGGTCTTTATCCCGCCCGCAGGAGTGGCGCCCATCATCGGAAGAGGGGAG 907  
 Db 407 ThrValHisGlnPheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGly 426  
 QY 908 CACATCAAAACAGCTCTCCCGGTTTCCAGCGCTCCATCAAGATTGACACCCGAAACA 967  
 Db 427 HisIleLysGlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAla 446  
 QY 968 CTGACTCCAAAGTTCGTATGTTATCATCTGACCGCCAGAGCGCCCAATTCAGGCT 1027  
 Db 447 ProAspAlaLysValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAla 466  
 QY 1028 CAGGGAAGAATCTATGGCAAACTCAAGAGGAGAACTTCTTTGGTCCCAAGGAGGAGTG 1087  
 Db 467 GlnGlyArgIleTyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluVal 486  
 QY 1088 AAGCTGGAGACCCACATACATGTCAGCATCAGCAGCTGGCGCGGCTCATTTGGCAAAGGT 1147

Db 487 LysLeuGluAlaHisIleA:gValProSerPheAlaAlaGlyArgValIleGlyLysGly 506  
Qy 1148 GGAARACGGTGAACGAGTTGCCAGAAATTCACGGCAGCTGAGGTGAGTACCAAGAGAC 1207  
Db 507 GlyLysThrValAsnGluLeuGlnAsnLeuSerAlaGluValValProArgAsp 526  
Qy 1208 CAGACCCCTGATGAGACGACGAGTGCATCGTGAATATCATCGACATTTCTATGCCAGT 1267  
Db 527 GlnThrProAspGluAsnAspGlnValValLysIleThrGlyHisPheTyrAlaCys 546  
Qy 1268 CAGATGCTCAACGAGATCCGAGACATCTCTGCCCCAGGTTAAGCAG---CAGCATCAG 1324  
Db 547 GlnValAlaGlnArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGln 566  
Qy 1325 AAGGGA---CAGAGTAACACGAGCCCGACGAGGAGGAG 1360  
Db 567 LysAlaLeuGlnSerGlyProGlnSerArgArgLys 579  
RESULT 6  
US-09-480-884A-176  
; Sequence 176, Application US/09480884A  
; Patent No. 6482597  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Hosken, Nancy A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Fanger, Gary R.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C6  
; CURRENT APPLICATION NUMBER: US/09/480,884A  
; CURRENT FILING DATE: 2001-08-27  
; NUMBER OF SEQ ID NOS: 330  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 176  
; LENGTH: 579  
; TYPE: PRF  
; ORGANISM: Homo sapiens  
US-09-480-884A-176  
Alignment Scores:  
Pred. No.: 7,3e-146 Length: 579  
Score: 1637.00 Matches: 336  
Percent Similarity: 83.66% Conservative: 43  
Best Local Similarity: 74.17% Mismatches: 58  
Query Match: 52.64% Indels: 16  
DB: 4 Gaps: 9  
US-09-270-437D-5 (1-1708) x US-09-480-884A-176 (1-579)  
Qy 32 CGGGAGCCATCATGAGCTGAATGGCCACAGTTCGAGAACCATGCCCTGAGGTCCTCC 91  
Db 133 ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla 152  
Qy 92 TACATCCCGATGACGATAGCA---CAGGGACCTGAGAATGGCGCCGAGGG-----142  
Db 153 TyrIleProAspGluMetAlaAlaGlnGlnAsnProLeuGlnProArgGlyArgArg 172  
Qy 143 GCCTTTGGCTCTCGGGTTCAGCCCGCAGGGCTCACCCTGTGCAGGGGGGCCAGGC 202  
Db 173 GlyLeuGlyGlnArgGlySerSerArgGlnGlySerPro-----GlySerValSer 189  
Qy 203 AAGCAGCAGCAAGTGGACATCCCTTCGGCTCTGTGTGCCACCCAGCATGTGGGTGCC 262  
Db 190 LysGlnLysProCysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAla 209  
Qy 263 ATTATTGGCAGGAGGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCCAGATA 322  
Db 210 IleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrLysIle 229  
Qy 323 GAGTGCATGAGGAGGAGACGAGGTGCAGCTGAAAGCCATCAGTGTGCACCTCCACC 382

Db 230 AspValHisArgLysGluAsnAlaGlyAlaAlaGlyLysSerIleThrIleLeuSerThr 249  
Qy 383 CCTGAGGGCTGCTCCTCGCTTGTAAAGATGATCTTGAGATTTATGCATATAAGAGGCTAAG 442  
Db 250 ProGluGlyThrSerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGln 269  
Qy 443 GACACCAAAACGGTGCAGAGGTTCCCTCGAAGATCTCTGGCCCATATAATACTTTGTAGGG 502  
Db 270 AspIleLysPheThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGly 289  
Qy 503 CGTCTCATTCGACGAGGAGGAGAACTGAGAGAGCTAGAGCAAGATACCGACACAAA 562  
Db 280 ArgLeuIleGlyLysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLys 309  
Qy 563 ATCACCATCTCCTCGTTGCAAGACCTTACCCCTTTACAAACCTCGAGAGGACCATCACTGTG 622  
Db 310 IleThrIleSerProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrVal 329  
Qy 623 AAGGGGCCATCGAGAAATTTGCGAGGCCGAGCAGGAATAATGAAGAAAGTTCCGGAG 682  
Db 330 LysGlyAsnValGluThrCysAlaLysAlaGluGluIleMethLysLysIleArgGlu 349  
Qy 683 GCCTATGAGAAATGATGTGGCTGCCATGAGC-----TCTCACCTGATCCCTGGCCTCAAC 736  
Db 350 SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn 369  
Qy 737 CTGGCTGCTGAGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCGC-----CCT 787  
Db 370 LeuAsnAlaLeuGlyLeuPheProThrSerGlyMetProProThrSerGlyPro 389  
Qy 788 CCCAGCAGCTTACTGGGGCTGCTCCCTATAGCTCTTTATGAGGCTCCGAGCAGCAG 847  
Db 390 ProSerAlaMetThr-----ProProTyrProGlnPheGluGlnSer---GluThrGlu 406  
Qy 848 ATGCTGCAAGTGTATTCGCCCGCAGCAGTGGGGCCCATCATCGCAAGAGGGCAG 907  
Db 407 ThrValHisGlnPheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGly 426  
Qy 908 CACATCAACAGCTCTCCCGGTTTCCAGCGCTCCATCAGATTGTCACACCCCGAACA 967  
Db 427 HisIleLysGlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAla 446  
Qy 968 CTGACTCCAAAGTTCTGTTATCATCATCTCGAGCGCCGAGAGGCCCAATTCAGGCT 1027  
Db 447 ProAspAlaLysValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAla 466  
Qy 1028 CAGGAGAAATCTATGGCAAACTCAAGAGGAGAACTTCTTGTGTCACAGGAGAGTG 1087  
Db 467 GlnGlyArgIleTyrGlyLysIleLysGluAsnPheValSerProLysGluGluVal 486  
Qy 1088 AAGCTGGAGACCCATACGTGTGCCAGCATCAGCAGCTGCCCGGGTCATTGGCAAGGT 1147  
Db 487 LysLeuGluAlaHisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGly 506  
Qy 1148 GGAATAACGGTGAACGAGTTTCAGAAATTCAGCGCAGCTGAGTGTGATACCAAGAGAC 1207  
Db 507 GlyLysThrValAsnGluLeuGlnAsnLeuSerAlaGluValValProArgAsp 526  
Qy 1208 CAGACCCCTGATGAGACGAGCAGCTCATCGTGAATAATCATCGACATTTCTATGCCAGT 1267  
Db 527 GlnThrProAspGluAsnAspGlnValValLysIleThrGlyHisPheTyrAlaCys 546  
Qy 1268 CAGATGCTCAACGAGATCCGAGACATCTCTGCCCCAGGTTAAGCAG---CAGCATCAG 1324  
Db 547 GlnValAlaGlnArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGln 566  
Qy 1325 AAGGGA---CAGAGTAACACGAGCCCGACGAGGAGGAG 1360  
Db 567 LysAlaLeuGlnSerGlyProGlnSerArgArgLys 579  
RESULT 7  
US-09-542-615A-176  
; Sequence 176, Application US/09542615A

```

; Patent No. 6518256
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C8
; CURRENT APPLICATION NUMBER: US/09/542.615A
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 176
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-542-615A-176

Alignment Scores:
Pred. No.: 7,3e-146 Length: 579
Score: 1637.00 Matches: 336
Percent Similarity: 83.66% Conservative: 43
Best Local Similarity: 74.17% Mismatches: 58
Query Match: 52.64% Indels: 16
DB: 4 Gaps: 9

US-09-270-437D-5 (1-1708) x US-09-542-615A-176 (1-579)
QY 32 CGGGAGCCATCATGAAGCTGAATGGCCACACAGTTGGAGAACCATGCCCTCGAAGGTCTCC 91
DB 133 ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla 152
QY 92 TACATCCCGATGACGAGATAGCA---CAGGACCTGAGAAATGGCGCCGAGGG----- 142
DB 153 TyrIleProAspGluMetAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArg 172
QY 143 GGCTTTGGCTCTCGGGTCACGCCCGCCAGGCTCACCTGTGGCAGGGGGGCCAGCC 202
DB 173 GlyLeuGlyGlnArgGlySerSerArgGlnGlySerPro-----GlySerValSer 189
QY 203 AAGCAGCAGGAAGTGGACATCCCTTCGGCTCCTCGGTGGCCGCCACCCAGTATGTGGTGCC 262
DB 190 LysGlnLysProCysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAla 209
QY 263 ATTTATGCGCAAGGAGGGGGCCACCATCGGCAACATCAAAACAGACCCAGTCAAGATA 322
DB 210 IleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIle 229
QY 323 GACGTGCATAGGAAGGAGACGACGAGTGCAGCTGAAAAGCCATCATGTGTGCATCCACC 382
DB 230 AspValHisArgLysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThr 249
QY 383 CCTCAGGCTCTCTCCGCTGTGATGATGATCTTGGAGATTATGCAATAAGAGGCTTAAG 442
DB 250 ProGluGlyThrSerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGln 269
QY 443 GACACCAAACGGCTGACGAGTCCCTCGAGATCTGGCCATATACCTTATCTTGTAGGG 502
DB 270 AspIleLysPheThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGly 289
QY 503 CGTCTCATTTGGCAAGGAGCGGAACCTCGAAGAGTAGACGAGATACCGAGACAAA 562
DB 290 ArgLeuIleGlyLysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLys 309
QY 563 ATCACCATCTCTGTTGCAAGACCTTACCTTTACACCCCTGAGAGACCATCACTGTG 622
DB 310 IleThrIleSerProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrVal 329
QY 623 AAGGGGGCCATCGAGAAATTGTTCAGGGCCGAGCAGGAAATATCAAGAAAGTTCCGGCAG 692

```

```

DB 330 LysGlyAsnValGluThrCysAlaLysAlaGluGluGluIleMetLysLysIleArgGlu 349
QY 683 GCCTATGAGATGATGTGGCTGCCATGAGC-----TCTCAGCTGATCCCTGGCTGAGC 736
DB 350 SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn 369
QY 737 CTGGCTGCTGTAGGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCCG-----CCT 787
DB 370 LeuAsnAlaLeuGlyLeuPheProThrSerGlyMetProProProThrSerGlyPro 389
QY 788 CCCAGCAGCGTTACTGGGGCTGCTCCCTATAGCTCTTTATGTCGAGCTCCCGAGCAGGAG 847
DB 390 ProSerAlaMetThr-----ProProTyrProGlnPheGluGlnSer---GluThrGlu 406
QY 848 ATGTGTCAGGTCTTTATCCCGCCCGCAGCAGTGGCGCCATCATCGGCAAGAGGGGAG 907
DB 407 ThrValHisGlnPheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGln 426
QY 908 CACATCAAAACAGCTCTCCCGGTTTCCAGCGCTCCATCAAGATTGACACCCCAACA 967
DB 427 HisIleLysGlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAla 446
QY 968 CCTGACTCCAAAGTTCGTATGTTATCATCTACTGGACCGCCAGAGGCCCAATTCAAGGCT 1027
DB 447 ProAspAlaLysValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAla 466
QY 1028 CAGGGAAGATCTATGCGAACTCAAGGAGGAGACTTCTTTGTCCTCCAAAGGAGGAAGT 1087
DB 467 GlnGlyArgIleTyrGlyLysIleLysGluGlnAsnPheValSerProLysGlnGluVal 486
QY 1088 AAGCTGGAGAGCCACATACGTGTGCCAGCATCAGCAGCTGGCGCGGTTCATTGGCAAAGT 1147
DB 487 LysLeuGluAlaHisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGly 506
QY 1148 GGAATAACGGTGAACAGATTGCGAATTTGACGGCAGCTGAGTGTGTAGTACCAAGAGAC 1207
DB 507 GlyLysThrValAsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAsp 526
QY 1208 CAGACCCCTGTAGAGAACGACGAGTCTCGTGAATCATCGGAATCATCGGACATTTCTATCCAGT 1267
DB 527 GlnThrProAspGluAsnAspGlnValValValLysIleThrGlyHisPheTyrAlaCys 546
QY 1268 CAGATGGCTCAACGGAAGATCCGACATCTCTGGCCCGCAGGTTAAGCAG---CAGCATCAG 1324
DB 547 GlnValAlaGlnArgLysIleGlnIleLeuThrGlnValLysGlnHisGlnGlnGln 566
QY 1325 AAGGGA---CAGAGTAACGAGCCCGCCAGCAGGAGGAG 1360
DB 567 LysAlaLeuGlnSerGlyProProGlnSerArgArgLys 579

RESULT 8
US-09-606-421B-176
; Sequence 176, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606.421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 176
; LENGTH: 579

```



```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-606-421B-176

Alignment Scores:
Pred. No.: 7,3e-146 Length: 579
Score: 1637.00 Matches: 336
Percent Similarity: 83.66% Conservative: 43
Best Local Similarity: 74.17% Mismatches: 58
Query Match: 52.64% Indels: 16
Dbs: 4 Gaps: 9

US-09-270-437D-5 (1-1708) x US-09-606-421B-176 (1-579)

QY 32 CGGGAGCCATCAAGCTGAATGGCCACAGTTGGAGAACCATGCGCTCGAAGGTCTCC 91
Db 133 ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla 152
QY 92 TACATCCCGCATGACGATAGCA---CAGGACCTGAGATGGGCCCGAGGG----- 142
Db 153 TyrIleProAspGluMetAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArg 172
QY 143 GCGTTTGCTCTCGGGGTCAGCCCGCCAGCGCTCACCTGTGGCAGCGGGGCCCGACCC 202
Db 173 GlyLeuGlyGlnArgGlySerSerArgGlnGlySerPro-----GlySerValSer 189
QY 203 AAGCAGCAGCAAGTGGACATCCCGCTTCGCTGCTGCTGCCACCCAGATATGTGGTGCC 262
Db 190 LysGlnLysProCysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAla 209
QY 263 ATTATTGCAAGGAGGGGCCACCATCGCAACATCACAACACAGACCCAGTCCAGATA 322
Db 210 IleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIle 229
QY 323 GACTGTGATAGGAAGGAGAACCGAGGTGCAGCTGAAAGAACCATCATGTGCACTCCACC 382
Db 230 AspValHisArgLysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThr 249
QY 383 CCTGAGGCTCTCTCGCTGTGAAGATGATCTTGAGATTTATGCTAAGAGGCTAAG 442
Db 250 ProGluGlyThrSerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGln 269
QY 443 GACACCAAAACGGCTGACGAGTTCCCTGAAATCCTGGCCCATATAACTTTGTAGGG 502
Db 270 AspIleLysPheThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGly 289
QY 503 GGTCTCATGCAAGGAGGACGGAACCTGAAGAGTAGACAGATACCGAGACAAA 562
Db 290 ArgLeuIleGlyLysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLys 309
QY 563 ATCAACCATCTCTCGTTGCAAGACTTACCTTTACAACTGAGAGGACCATCACTGTG 622
Db 310 IleThrIleSerProLeuGlnGluLeuThrLeuTyAsnProGluArgThrIleThrVal 329
QY 623 AAGGGGCCATTCAGAAATTTGTCAGGCGCAGCAGAAATAATGAAGAAAGTTCGGAG 682
Db 330 LysGlyAsnValGluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGlu 349
QY 683 GCCTATGAGATGATGGCTGCCATGAGC-----TCTCACCTGATCCCTGGCCCTGAAC 736
Db 350 SerTyrgluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn 369
QY 737 CTGGCTCTGTAGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCG-----CCT 787
Db 370 LeuAsnAlaLeuGlyLeuPheProProThrSerGlyMetProProThrSerGlyPro 389
QY 788 CCCAGCAGCGTTACTGGGGTGCTCCCTATAGCTCTTTATGAGCTCCCGAGCAGGAG 847
Db 390 ProSerAlaMetThr-----ProProTyProGlnPheGluGlnSer---GluThrGlu 406
QY 848 ATGTGTCAGGTGTTTATCCCGCCCGCAGCAGTGGGCCCATCATCGCAAGAGGGGCGAG 907
Db 407 ThrValHisGlnPheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGln 426

```

```

QY 908 CACATCAAAACAGCTCTCCGGTTTCCAGCGCTCCATCAAGATTGCACCCCGAAACA 967
Db 427 HisIleLysGlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAla 446
QY 968 CTGACTCCAAAGTTCTGATGTTATCATCTACTGACCGCCAGCGGCCCAATTCAAGCT 1027
Db 447 ProAspAlaLysValArgMetValIleIleThrGlyProGluAlaGlnPheLysAla 466
QY 1028 CAGGGAAGATCTATGGCAAACTCAAGGAGGAGACTTCTTTGGTCCCAAGGAGGAATG 1087
Db 467 GlnGlyArgIleTyGlyLysIleLysGluAsnPheValSerProLysGlnGluVal 486
QY 1088 AAGCTGGAGACCCACATACGTGTGCCAGCAGTCAGCAGCTGCGCGGTTCATTGGCAAAAGT 1147
Db 487 LysLeuGluAlaHisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGly 506
QY 1148 GCAAAAACGGTGAACGAGCTTGCAGATTTCGACGAGCTGAGTGGTAGTACCAAGAC 1207
Db 507 GlyLysThrValAsnGluLeuGlnAsnLeuSerAlaGluValValProArgAsp 526
QY 1208 CAGACCCCTGATGAGAACAGACCCAGGTCTATCGTGAATAATCATCGGACATTTCTATGCCAGT 1267
Db 527 GlnThrProAspGluAsnAspGlnValValLysIleThrGlyHisPheTyAlaCys 546
QY 1268 CAGATGGCTCAACGGAAGATCCGAGACATCTCTGGCCCGCAGGTTAAGCAG---CAGCATCAG 1324
Db 547 GlnValAlaGlnArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGln 566
QY 1325 AAGGGA---CAGAGTAAACGAGGCCCGCCAGGACCGAGGAAG 1360
Db 567 LysAlaLeuGlnSerGlyProGlnSerArgArgLys 579

RESULT 9
US-08-021-608D-10
; Sequence 10, Application US/08021608D
; Patent No. 5580760
; GENERAL INFORMATION:
; APPLICANT: LEVENS, DAVID L., DUNCAN,
; APPLICANT: ROBERT C., AND AVIGAN, MARK I.
; TITLE OF INVENTION: NOVEL FUSE BINDING
; TITLE OF INVENTION: PROTEIN AND CDNA THEREFOR
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/021,608D
; FILING DATE: 22-FEB-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: WILLIAM S. FEILER
; REGISTRATION NUMBER: 26,728
; REFERENCE/DOCKET NUMBER: 2026-4063
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 643

```



```

; TYPE: Amino Acid
; STRANDEDNESS: Single
; TOPOLOGY: Unknown
; MOLECULE TYPE: Peptide/Protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Human
; CELL LINE: HL60
; FEATURE:
; OTHER INFORMATION:
; OTHER INFORMATION: Amino Acid 148 (Xaa) is Met or Ile
US-08-021-608D-10

Alignment Scores:
Pred. No.: 1,15e-14 Length: 643
Score: 248.50 Matches: 140
Percent Similarity: 37.80% Conservatives: 97
Best Local Similarity: 22.33% Mismatches: 202
Query Match: 7.99% Indels: 189
DB: 1 Gaps: 32

```

US-09-270-437D-5 (1-1708) x US-08-021-608D-10 (1-643)

```

QY 98 CCGGAT--GAGCAGATAGCAGGGACCTGAGAATGGCGCGCAGGGGGCTTTGGCTCT 154
Db 73 ProAspAlaLysLysValAla-----ProGlnAsn-----AspSerPheGlyThr 87
QY 155 CGGGGTGACCCCGCCAGGGCTCACCTGTGCGAGGGGGGCCCGACCGACGACGACAA 214
Db 88 GlnLeuProProMetHis-----GlnGlnGlnArg 97
QY 215 GTGGACATCCCTTCGGCTCTGTCGCCACCCAGTCATGTGGTGCCATTATTGGCAAG 274
Db 98 SerValMetThrGluGluTyrlsValProAspGlyMetValGlyPheIleIleGlyArg 117
QY 275 GAGGGGGCCACATCCGCAACATCACAAACAGACCCAGTCACAGATAGACGTGCATAGG 334
Db 118 GlyGlyGluGlnIleSerArgIleGlnGlnIleGlnSerGlyCysIleGlnIle---Ala 136
QY 335 AAGGAGAACGCGAGTGCAGCTGAAAGCCATCATGTGTCACCTCCACCCCTGAGGGGTGC 394
Db 137 ProAspSerGlyGlyLeuProGluArgSerCys***LeuThrGlyThrProGluSerVal 156
QY 395 TCCTCCGCTTGAAGATCTTGAGATTATGATATAA-----GAGGCTTAAGACACC 448
Db 157 GlnSerAlaLysArgLeuLeuAspGlnIleValGluLysGlyArgProAlaProGlyPhe 176
QY 449 AAAACGGGTGAC-----GAGGTTCCCTGAGATCTCTGGGCCCATATAACTTTGTA 499
Db 177 HisHisGlyAspGlyProGlyAsnAlaValGlnGluIleMetIleProAlaSerLysAla 196
QY 500 GGGCGCTCTCATTGGCAAGAGGAGCGAACTGAAGAGGTAGACCAAGATACCGAGACA 559
Db 197 GlyLeuValIleGlyLysGlyGlyGluThrIleLysGlnLeuGlnGluArgAlaGlyVal 216
QY 560 AAATCACCATCTCTCTTCGACAGACCTTACCTTTACACCCCTGAGAGGACC----- 613
Db 217 LysMetValMet-----IleGlnAsp-----GlyProGlnAsnThrGlyAla 230
QY 614 -----ATCACTGTGAAGGGGCCCATCGAGAATTGTCAGGGCGGAGCAGCAATA 664
Db 231 AspLysProLeuArgIleThrGlyAspProTyrlsValGlnGlnAlaLysGluMetVal 250
QY 665 ATGAAGAAAGTTCGGGAG-----GCCTATGAGAATGATGTGGTGGCCCATG 709
Db 251 LeuGluLeuIleArgAspGlnGlyGlyPheArgGluValAsnGluTyrlsGly----- 268
QY 710 AGCTCTCACTGATCCCTGGCCCTGACCTGCTGCTAGCTCTTTCCCGACGTTTCATCC 769
Db 268 -----
QY 770 AGCGCAGTCCGCGCGCTCCGAGCAGCGTTACTGGGGTGTCTCCCTATAGCTCTTTATG 829
Db 770 -----

```

```

Db 269 -----SerArgIleGlyGly----- 273
QY 830 CAGGCTCCCGAGCAGGAGATGGTGCCAGTGTTCATCCCGCCCGCCAGGAGTGGCGGCATC 889
Db 274 -----AsnGluGlyIleAspValProIleProArgPheAlaValGlyIleVal 289
QY 890 ATCGGCAAGAGGGGCGAGCAGCATCAACAGCTCTCCCGGTTCGCCAGCGCTCCATCAAG 949
Db 290 IleGlyArgAsnGlyGluMetIleLysLysIleGlnAsnAspAlaGlyValArgIleGln 309
QY 950 ATTGCACACCCCGAA-----ACACCTGACTCCAAAGTTCTGATGTATTCATCACTGGA 1003
Db 310 PheLysProAspAspGlyThrThrProGlu-----ArgIleAlaGlnIleThrGly 326
QY 1004 CGCCGAGAG---GCCCAATTCAAGGCTCAG----- 1030
Db 327 ProProAspArgCysGlnHisAlaAlaGluIleThrAspLeuLeuArgSerValGln 346
QY 1031 -----GGAGGAATCTATGCGCAACTATGCGCAACTCAAGGAGGAG 1060
Db 347 AlaGlyAsnProGlyGlyProGlyProGlyGlyArgGlyArgGlyGlyGlnGlyAsn 366
QY 1061 AACTTCTTTGTCCTCCCAAGGAGGAAGTGAAGCTGGAGACCCACATAGCTGTGCCGACATCA 1120
Db 367 TrpAsnMetGlyProProGlyGlyLeuGln---GluPheAsnPheIleValProThrGly 385
QY 1121 GCAGCTGGCGGGTCAATTGGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1180
Db 386 LysThrGlyLeuIleIleGlyLysGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 405
QY 1181 GCAGCTGAGTGTGTGTACAGAGACACGACCCCTGTGATGAGAACACACCGAGTGC---ATC 1237
Db 406 GlyAlaArgIleGluLeuGlnArgAsnProProProAsnAlaAspProAsnMetLysLeu 425
QY 1238 GTGAAAATCATCGGACAT-----TTCTATGCCAGTCAGATGGCTCAACGGAAG 1285
Db 426 PheThrIleArgGlyThrProGlnGlnIleAspTyrlsAlaArgGlnLeuIleGluLys 445
QY 1286 AT-----CCGAGACATCT---GSCCCAGGTGAAGCAGCAGCAGCAGCAGCAGCAG 1335
Db 446 IleGlyGlyProValAsnProLeuGlyProProValProHisGlyProHisGlyVal--- 464
QY 1336 TAACAGGCGCCAGGACGGA-----GGAAGTACCAGCCCTC----- 1373
Db 465 ---ProGlyPro--HisGlyProProGlyProProGlyProGlyThrProMetGlyProT 483
QY 1374 ---CCTGCTCCCTTNGAGTCCAGGA----- 1394
Db 483 TyrAsnProAlaProTyrlsAsnProGlyProGlyProAlaProHisGlyProProAlaP 503
QY 1395 -----CAACAAACGGGCGAGAAATCG 1413
Db 503 roTyrlAlaProGlnGlyTrpGlyAsnAlaTyrlsProHisTrpGlnGlnAlaProProA 523
QY 1414 AGAGTGTGTCTCCCGCAGCGCTGAGAAATGAGTGGGAATCCGGGACACACNTGGCGCGGG 1473
Db 523 sp-----ProAlaLysAlaGlyThrAspProAsnSerAla-AlaTrpAlaAla 538
QY 1474 CTGTAGATCAGGTTTCGCCCTTGTGAGAAAGATGTTCCAGTGAGGAACCTGATCTN 1533
Db 539 Tyr-----TyrlAlaHisTyrlsGlnGlnGlnAla----- 548
QY 1534 TCAGCCGCCAAACACCCACCCCAATTGGC---CCAACTGTNTGCCCTCGGGGTGTCAGA 1590
Db 549 GlnProProAlaAlaProAlaGlyAlaProThrThrThrThrGlnThrAsnGlyGlnGly 568
QY 1591 AATTNTAGCGCAGGACACTTTTAAACGTGATGTTTAAAGAGCTCTCCAGGCCGCCACC 1650
Db 569 Asp-----GlnGlnAsnProAlaProAla 576
QY 1651 AAGAGGGTGGATCACACC-----TCAGTGGGAAGAGAAATA 1686
Db 577 GlyGlnValAspTyrlsThrLysAlaTrpGluGluTyrlsLysLysMetGlyGlnAlaVal 596

```





QY 449 AAAACGGCTGAC-----GAGTTCCTCCTGAAGATCTCGGCCCAATAATACTTTGTA 499  
 Db 177 HisHisGlyAspGlyProGlyAsnAlaValGlnGluMetIleProAlaSerLysAla 196  
 QY 500 GGGCTCTCATTGGCAAGAGGAGCGAACTGAAGAGGTAGAGCAAGATACCGAGACA 559  
 Db 197 GlyLeuValIleGlyLysGlyGlyGluThrIleLysGlnLeuGlnAlaGlyVal 216  
 QY 560 AAATCACCCTCTCTCTGTCAGACCTTACCTTTACAACTCTGAGAGACC-----613  
 Db 217 LysMetValMet-----IleGlnAsp-----GlyProGlnAsnThrGlyVala 230  
 QY 614 -----ATCACTGTGAAGGGCCCATCGAGAATTTGTTCAGGGCCGAGCAGAAATA 664  
 Db 231 AsplysProLeuArgIleThrGlyAspProTyrLysValGlnGlnAlaLysGluMetVal 250  
 QY 665 ATGAAGAAGTTCGGAG-----GCCTATGAGAAATGATGTGGTGGCCATG 709  
 Db 251 LeuGluLeuIleArgAspGlnGlyGlyPheArgGluValArgAsnGluTyrGly-----268  
 QY 710 AGCTCTACCTGATCCCTGGCCTGAACCTGGCTCTGTAGTCTTTTCCCGAGCTTCATCC 769  
 Db 268 -----268  
 QY 770 AGCGAGTCCCGCGCCTCCAGCAGCGTACTGGGCTGTCTCCCTATAGCTCCTTTATG 829  
 Db 269 -----SerArgIleGly-----273  
 QY 830 CAGGCTCCGAGCAGAGATGGTGCAGGTGTTTATCCCGCCAGGAGTGGGGCCATC 889  
 Db 274 -----AsnGluGlyIleAspValProIleProArgPheAlaValGlyIleVal 289  
 QY 890 ATCGGAAGAAGGGCAGACATCAACAGCTCTCCCGTTTGGCAGCGCTCCATCAAG 949  
 Db 290 IleGlyArgAsnGlyGluMetIleLysIleGlnAsnAspAlaGlyValArgIleGln 309  
 QY 950 ATTGCACACCCGAA-----ACACCTGACTCCAAAGTTGCTATGGTTATCATCCTGGA 1003  
 Db 310 PheLysProAspAspGlyThrThrProGlu-----ArgIleAlaGlnIleThrGly 326  
 QY 1004 CCGCCAGAG-----GCCCAATTCAGGCTCAG-----1030  
 Db 327 ProProAspArgCysGlnHisAlaGluIleThrAspLeuLeuArgSerValGln 346  
 QY 1031 -----GGAGAATCTATGGCAAACTCAAGAGGAG 1060  
 Db 347 AlaGlyAsnProGlyGlyProGlyArgGlyArgGlyArgGlyGlnGlyAsn 366  
 QY 1061 AACTTCTTTGTTCCCAAGGAGGAAGTGAAGTGGAGACCCACATACGTGTGCCAGCATCA 1120  
 Db 367 TrpAsnMetGlyProProGlyGlyLeuGln-----GluPheAsnPheIleValProThrGly 385  
 QY 1121 CGAGTGGCCGGGTCAATTGGCAAGGGTGGAAAAACGGTGAACGAGTTGCAGAATTTGACG 1180  
 Db 386 LysThrGlyLeuIleIleGlyLysGlyGlyGluThrIleLysSerIleSerGlnSer 405  
 QY 1181 CGAGTGGGTGGTAGTACCAAGAGACAGACCCCTGATGAGAACGACGAGTCT---ATC 1237  
 Db 406 GlyAlaArgIleGluLeuGlnArgAsnProProProAsnAlaAspProAsnMetLysLeu 425  
 QY 1238 GTGAAATCATCGGACAT-----TTCTATGCCAGTCAAGTGGTCAACGGAAG 1285  
 Db 426 PheThrIleArgGlyThrProGlnGlnIleAspTyrAlaArgGlnLeuIleGluGlyLys 445  
 QY 1286 AT-----CCGAGACATCCT---GGCCAGGTTTAAGCAGCAGCATCAGAGAGCAGAG 1335  
 Db 446 IleGlyGlyProValAsnProLeuGlyProProValProHisGlyProHisGlyVal---464  
 QY 1336 TAACAGGCCCCAGCAGCGA-----GGAAAGTGACAGCCCTC-----1373  
 Db 465 ---ProGlyPro---HisGlyProProGlyProGlyProGlyThrProMetGlyProT 483  
 QY 1374 -----CCTGTCCCTTNGAGTCCAGGA-----1394

Db 483 y-AsnProAlaProTyrAsnProGlyProProGlyProAlaProHisGlyProProAlaP 503  
 QY 1395 -----CAACAACGGCGAGAAATCG 1413  
 Db 503 roTyrAlaProGlnGlyTrpGlyAsnAlaTyr-ProHisTrpGlnGlnAlaProProA 523  
 QY 1414 AGAGTGTGTCTCCCGGAGCCCTGAGATGAGTGGGAATCCGGGACACTGGGCCGGG 1473  
 Db 523 sp-----ProAlaLysAlaGlyThrAspProAsnSerAla-AlaTrpAlaAla 538  
 QY 1474 CTGTAGATCAGGTTTGGCCACTTGATGAGAAAGATGTTCAGTGAAGAACCTGATCTN 1533  
 Db 539 Tyr-----TyrAlaHisIleTyrTrpGlnGlnAla-----548  
 QY 1534 TCAGCCCAACACCCACCAATTCGC---CAACACTGTNTGCCCTCGGGGTGTGAGA 1590  
 Db 549 GlnProProAlaAlaProAlaGlyAlaProThrThrThrGlnThrAsnGlyGlnGly 568  
 QY 1591 AATTNTAGCGCAAGGCACCTTTTAAAGAGCTCTCCAGGCCCCACC 1650  
 Db 569 Asp-----GlnGlnAsnProAlaProAla 576  
 QY 1651 AAGAGGTGATCAACACC-----TCAGTGGGAAGAAAAATA 1686  
 Db 577 GlyGlnValAspTyrThrLysAlaTrpGluGluTyrTyrLysLysMetGlyGlnAlaVal 596  
 QY 1687 AATTTCCTTCAGGT 1701  
 Db 597 ProAlaProThrGly 601  
 RESULT 12  
 US-09-261-855-18  
 ; Sequence 18, Application US/09261855A  
 ; Patent No. 6255055  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ross, Jeffrey  
 ; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN  
 ; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE  
 ; FILE REFERENCE: 960296.95131  
 ; CURRENT APPLICATION NUMBER: US/09/261,855A  
 ; CURRENT FILING DATE: 1999-03-03  
 ; NUMBER OF SEQ ID NOS: 46  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 18  
 ; LENGTH: 49  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus  
 US-09-261-855-18  
 Alignment Scores:  
 Pred. No.: 5,86e-15 Length: 49  
 Score: 245.00 Matches: 49  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 7.88% Indels: 0  
 DB: 3 Gaps: 0  
 US-09-270-437D-5 (1-1708) x US-09-261-855-18 (1-49)  
 QY 476 ATCTCGGCCCAATAATACTTTGTAGGGCTCTCATTTGCAAGGAGGACGGAACCTGAAG 535  
 Db 1 IleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGlyArgAsnLeuLys 20  
 QY 536 AAGGTAGAGCAAGATACCGAGCAAAAATACCATCTCTCTCTGTCGAGAGACTTACCTT 595  
 Db 21 LysValGluGlnAspThrGluThrLysIleThrIleSerSerLeuGlnAspLeuThrLeu 40  
 QY 596 TACAACCTCTGAGAGACCATCACTGTG 622  
 Db 41 TyrAsnProGluArgThrIleThrVal 49  
 RESULT 13

US-08-021-608D-2  
; Sequence 2, Application US/08021608D  
; Patent No. 5580760  
; GENERAL INFORMATION:  
; APPLICANT: LEVENS, DAVID L., DUNCAN  
; APPLICANT: ROBERT C., AND AVIGAN, MARK I.  
; TITLE OF INVENTION: NOVEL FUSE BINDING  
; TITLE OF INVENTION: PROTEIN AND CDNA THEREFOR  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN & FINNEGAN  
; STREET: 345 PARK AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY DISK  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WORDPERFECT 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/021,608D  
; FILING DATE: 22-FEB-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WILLIAM S. FEILER  
; REGISTRATION NUMBER: 26,728  
; REFERENCE/DOCKET NUMBER: 2026-4063  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 758-4800  
; TELEFAX: (212) 751-6849  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 644  
; TYPE: Amino Acid  
; STRANDEDNESS: Single  
; TOPOLOGY: Unknown  
; MOLECULE TYPE: Peptide/Protein  
; HYPOTHETICAL: No  
; ORIGINAL SOURCE:  
; ORGANISM: Human  
; CELL LINE: HL60  
; FEATURE:  
; OTHER INFORMATION: Amino Acid 149 (Xaa) is Met or Ile  
US-08-021-608D-2  
Alignment Scores:  
Pred. No.: 2,468-14 Length: 644  
Score: 245.00 Matches: 148  
Percent Similarity: 36.46% Conservative: 93  
Best Local Similarity: 22.39% Mismatches: 219  
Query Match: 7.88% Indels: 202  
DB: 1 Gaps: 33  
US-09-270-437D-5 (1-1708) x US-08-021-608D-2 (1-644)  
QY 71 AACATGCCCTGAAGTCTCTACATCCCGATGACGAGATAGCACAG-----GGACCT 124  
Db 28 AsnAspAlaPheLysAspAlaLeuGlnArgAlaArgGlnIleAlaAlaLysIleGly 47  
QY 125 GGAATGGG-----CGCCGAGGGGCTTGGCTCTCGGCTCAGCCCGCCAGGGC 175  
Db 48 AspAlaGlyThrSerLeuAsnSerAsnAspTyrGlyTyrGlyGlnLysArg----- 65  
QY 176 TCACCTGTGGCAGCGGG----- 193  
Db 66 ---ProLeuGluAspGlyAspGlnProAspAlaLysLysValAlaProGlnAsnAspSer 84

QY 194 -----GCCCGAGCCAGCAGCAGCAGTGGACATCCCTTT-----CG 232  
Db 85 PheGlyThrGlnLeuProMetHisGlnGlnSerArgSerValMetThrGluGlu 104  
QY 233 CTCCTGGTGGCCACCATGATGCGGTGCATATTGGCAAGAGGGGGCCACCTCCGC 292  
Db 105 TyrLysValProAspGlyMetValGlyPheIleIleGlyArgGlyGluGlnIleSer 124  
QY 293 AACATCAAAAACAGACACCCAGTCCCAAGATAGAGTGCATAGGAAGGAGAACCCAGTGCA 352  
Db 125 ArgIleGlnGlnGluSerGlyCysLysIleGlnIle---AlaProAspSerGlyGlyLeu 143  
QY 353 GCTGAAAAGCCATCAGTGTGCTACCTCCACCTGAGGGCTGCTCTCCCTGTGTAAGATG 412  
Db 144 ProGluArgSerCys\*\*\*LeuThrGlyThrProGluSerValGlnSerAlaLysArgLeu 163  
QY 413 ATCTTGGAGATTATGATAA-----GAGGTAAAGCACACCAAAACCGCTCAC----- 460  
Db 164 LeuAspGlnIleValGlnLysGlyArgProAlaProGlyPheHisGlyAspGlyPro 183  
QY 461 ---GAGGTTCCCTGAAGATCCTGCCCATATAACTTTGTAGGGGCTCTCATTTGCAAG 517  
Db 184 GlyAsnAlaValGlnGluIleMetIleProAlaSerLysAlaGlyLeuValIleGlyLys 203  
QY 518 GAAGGACGGACCTGAAGAGTGTAGAGCAGATACCGACACAAAATACCATCTCTCTCG 577  
Db 204 GlyGlyGluThrIleLysGlnLeuGlnGluArgAlaGlyValLysMetValMet----- 221  
QY 578 TTGCAAGACCTTACCTTTACAACTCTGAGAGACC-----ATCACTGTG 622  
Db 222 IleGlnAsp-----GlyProGlnAsnThrGlyAlaAspLysProLeuArgIle 237  
QY 623 AAGGGGCCCATCGAATTTGTCAGGGCGGAGCAGGAAATATGAGAAAGTTCCGGAG 682  
Db 238 ThrGlyAspProTyrLysValGlnGlnAlaLysGluMetValLeuGluLeuIleArgSp 257  
QY 683 -----GCCTATGAGAAATGATGGCTGCCATGAGCTCTCACCTGATCCCT 727  
Db 258 GlnGlyGlyPheArgGluValArgAsnGluTyrGly----- 269  
QY 728 GGCTGAACCTGGCTGCTGTAGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCGCCCT 787  
Db 269 ----- 269  
QY 788 CCCAGCAGCTTACTGGGGTGTCTCCCTATAGTCTCTTTATGCAGGCTCCCGAGCAGGAG 847  
Db 270 ---SerArgIleGlyGly-----AsnGlu 276  
QY 848 ATGTCAGCTGTTTATCCCGCCCGCAGCAGTGGCGCCATCATCGCAAGAGGGCAG 907  
Db 277 GlyIleAspValProIleProArgPheAlaValGlyIleValIleGlyArgAsnGlyGlu 296  
QY 908 CACATCAAAACAGCTCTCCCGGTTTGGCAGCGGCTCCATCAAGATTGCACACCCGAA--- 964  
Db 297 MetIleLysIleGlnAsnAspAlaGlyValArgIleGlnPheLysProAspGly 316  
QY 965 ---ACACCTGACTCCAAAGTTCGTATGTTATCATCTGAGCCGCGCAGAG---GCCCAA 1018  
Db 317 ThrThrProGlu-----ArgIleAlaGlnIleThrGlyProProAspArgCysGln 333  
QY 1019 TTCAAGGCTCAG----- 1030  
Db 334 HisAlaAlaGluIleIleThrAspLeuLeuArgSerValGlnAlaGlyAsnProGlyGly 353  
QY 1031 -----GGAAGATCTATGCAAACTCAAGGAGGAGAACTTCTTTGCTCCCAAG 1078  
Db 354 ProGlyProGlyArgGlyArgGlyArgGlyGlnGlyAsnTrpAsnMetGlyProPro 373  
QY 1079 GAGGAAGTGAAGCTGGAGACCCACATACGTGTGCCAGCATCAGCAGCTGCCCGGGTCATT 1138  
Db 374 GlyGlyLeuGln---GluPheAsnPheIleValProThrGlyLysThrGlyLeuIleIle 392









```
QY 1609 TTTTAAAGCTGGATTGTTTAAAGAAGCTCTCAGGCCCCACCACCAAGAGGGTGGATCAGACC 1668
Db 571 ::::: |||
GlnGlnAsnProAlaProAlaGlyGlnValAspTyrThr 583
QY 1669 -----TCAGTGGGAAGAAAAATAAAATTTCCTTCAGGT 1701
Db 584 LysAlaTrpGluGluTyrTyrLysLysMetGlyGlnAlaValProAlaProThrGly 602
```

Search completed: July 13, 2004, 12:14:43  
Job time : 47 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 13, 2004, 17:54:49 ; Search time 134 Seconds  
(without alignments)  
7073.553 Million cell updates/sec

Title: US-09-270-437D-5  
Perfect score: 1708  
Sequence: 1 aggaagctgcgcagccgc.....atttccttcagggttttaaaa 1708

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Query Match | Length | ID                 | Description        |
|------------|--------|-------------|--------|--------------------|--------------------|
| 1          | 1703   | 99.7        | 1708   | US-09-061-709-5    | Sequence 5, Appli  |
| 2          | 1703   | 99.7        | 1708   | US-09-899-651-5    | Sequence 5, Appli  |
| 3          | 1669.2 | 97.7        | 1946   | US-09-061-709-7    | Sequence 7, Appli  |
| 4          | 1669.2 | 97.7        | 1946   | US-09-899-651-7    | Sequence 7, Appli  |
| 5          | 1165   | 68.2        | 2224   | US-09-261-855-1    | Sequence 1, Appli  |
| 6          | 582.2  | 34.1        | 1740   | US-09-643-597-347  | Sequence 347, App  |
| 7          | 582.2  | 34.1        | 1740   | US-09-542-615A-347 | Sequence 347, App  |
| 8          | 582.2  | 34.1        | 1740   | US-09-606-421B-347 | Sequence 347, App  |
| 9          | 580.6  | 34.0        | 4159   | US-09-061-709-4    | Sequence 4, Appli  |
| 10         | 580.6  | 34.0        | 4159   | US-09-899-651-4    | Sequence 4, Appli  |
| 11         | 580.6  | 34.0        | 4181   | US-09-643-597-175  | Sequence 175, App  |
| 12         | 580.6  | 34.0        | 4181   | US-09-480-884A-175 | Sequence 175, App  |
| 13         | 580.6  | 34.0        | 4181   | US-09-542-615A-175 | Sequence 175, App  |
| 14         | 580.6  | 34.0        | 4181   | US-09-606-421B-175 | Sequence 175, App  |
| 15         | 504    | 29.5        | 3283   | US-09-061-709-8    | Sequence 8, Appli  |
| 16         | 504    | 29.5        | 3283   | US-09-899-651-8    | Sequence 8, Appli  |
| 17         | 489.2  | 28.6        | 3412   | US-09-061-709-6    | Sequence 6, Appli  |
| 18         | 489.2  | 28.6        | 3412   | US-09-899-651-6    | Sequence 6, Appli  |
| 19         | 53.6   | 3.1         | 7218   | US-08-232-463-14   | Sequence 14, Appli |
| 20         | 45.6   | 2.7         | 598    | US-09-669-751-39   | Sequence 39, Appli |
| 21         | 40.6   | 2.4         | 2850   | US-08-503-172-5    | Sequence 5, Appli  |
| 22         | 40.6   | 2.4         | 2850   | US-09-135-211-5    | Sequence 5, Appli  |
| 23         | 39.6   | 2.3         | 1982   | US-09-016-434-1067 | Sequence 1067, App |
| 24         | 39.6   | 2.3         | 1982   | US-09-825-497A-40  | Sequence 40, Appli |
| 25         | 39.6   | 2.3         | 1987   | US-09-825-497A-39  | Sequence 39, Appli |
| 26         | 39.6   | 2.3         | 1987   | US-09-517-779-1    | Sequence 1, Appli  |
| 27         | 37.6   | 2.2         | 1896   | US-09-343-011B-4   | Sequence 4, Appli  |

#### ALIGNMENTS

##### RESULT 1

US-09-061-709-5  
; Sequence 5, Application US/09061709B  
; Patent No. 6297364  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Yao-Tseng  
; APPLICANT: Gure, Ali  
; APPLICANT: Tsang, Solam  
; APPLICANT: Stockert, Elisabeth  
; APPLICANT: Jager, Elke  
; APPLICANT: Knuth, Alexander  
; APPLICANT: Old, Lloyd J.  
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated  
; FILE REFERENCE: LUD 5538  
; CURRENT APPLICATION NUMBER: US/09/061,709B  
; CURRENT FILING DATE: 1998-04-17  
; NUMBER OF SEQ ID NOS: 8  
; SEQ ID NO 5  
; LENGTH: 1708  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
US-09-061-709-5

Query Match 99.7%; Score 1703; DB 3; Length 1708;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGAGCTGCGCGACCGCCCGAGTTTACCCGGGAGCCATCATGAAGCTGAATGGCCA 60  
Db 1 AGGGAGCTGCGCGACCGCCCGAGTTTACCCGGGAGCCATCATGAAGCTGAATGGCCA 60

Qy 61 CCAGTTGGAGAACCATGSCCTTGAAGTCTCTTACATCCCGATGAGCAGATGACAGGG 120  
Db 61 CCAGTTGGAGAACCATGSCCTTGAAGTCTCTTACATCCCGATGAGCAGATGACAGGG 120

Qy 121 ACCTGAGATGGCGCGCGAGGGGCTTTGGCTCTCGGGTTCAGCCCGCGAGGGCTCACC 180  
Db 121 ACCTGAGATGGCGCGCGAGGGGCTTTGGCTCTCGGGTTCAGCCCGCGAGGGCTCACC 180

Qy 181 TGTGGCAGGGGGGGGGCCAGCAAGCAGCAGAGTGGACATCCCGCTTCGGCTCCTGGT 240  
Db 181 TGTGGCAGGGGGGGGGCCAGCAAGCAGCAGAGTGGACATCCCGCTTCGGCTCCTGGT 240

Qy 241 GCCCACCAGTAGTGTGGGTGGCCATTATTTGGCAAGAGGGGGGCCACCATCCGCAACATCAC 300  
Db 241 GCCCACCAGTAGTGTGGGTGGCCATTATTTGGCAAGAGGGGGGCCACCATCCGCAACATCAC 300

Qy 301 AAAACAGACCCAGTCCCAAGATAGACGTGCTATAGGAGGAGAACGCAGGTGCAGCTGAAA 360



Db 241 GCCACCCAGTATGTGGGTGCCATTATTGGCAAGAGGGGGCCACCATTCCGCAACATCAC 300  
Qy 301 AAAACAGACCCAGTCCAAAGATAGACGTGTGCATAGGAAGGAGAACCGAGGTGCAGCTGAAAA 360  
Db 301 AAAACAGACCCAGTCCAAAGATAGACGTGCATAGGAAGGAGAACCGAGGTGCAGCTGAAAA 360  
Qy 361 AGCCATCAGTGTGACTCCACCCCTGAGGGTGTCTCCCTCCGTTGTAAAGATGATCTTGGGA 420  
Db 361 AGCCATCAGTGTGACTCCACCCCTGAGGGTGTCTCCCTCCGTTGTAAAGATGATCTTGGGA 420  
Qy 421 GATTATGATCAAGAGGCTGAAGGACACCAAAACGGCTGACGAGGTTCCCTCTGAAGATCCT 480  
Db 421 GATTATGATCAAGAGGCTTAGGACACCAAAACGGCTGACGAGGTTCCCTCTGAAGATCCT 480  
Qy 481 GSCCATTAATACTTTGTAGGGCGTCTCATTTGGCAAGGAAGGAACCGAACCTTGAAGAGGT 540  
Db 481 GSCCATTAATACTTTGTAGGGCGTCTCATTTGGCAAGGAAGGAACCGAACCTTGAAGAGGT 540  
Qy 541 AGAGCAAGATACCGAGACAAAAATCACCATCTCTCTGTTCAAGACCTTACCTTTTCAA 600  
Db 541 AGAGCAAGATACCGAGACAAAAATCACCATCTCTCTGTTCAAGACCTTACCTTTTCAA 600  
Qy 601 CCCTGAGAGGACCACTACTGTGAAGGGGCCATCGAGAAATTTGTGACGGGCCGAGCAGGA 660  
Db 601 CCCTGAGAGGACCACTACTGTGAAGGGGCCATCGAGAAATTTGTGACGGGCCGAGCAGGA 660  
Qy 661 AATATGAGAAAGTTCGGAGGCTATGAGAAATGATGTGCTGCCATGAGCTCTCACCT 720  
Db 661 AATATGAGAAAGTTCGGAGGCTATGAGAAATGATGTGCTGCCATGAGCTCTCACCT 720  
Qy 721 GATCCCTGGCTGAACCTGGTGTGTAGGTCTTTTCCAGCTTCATCCAGCGCAGTCCC 780  
Db 721 GATCCCTGGCTGAACCTGGTGTGTAGGTCTTTTCCAGCTTCATCCAGCGCAGTCCC 780  
Qy 781 GCGGCTCCAGCAGCGTTACTGGGGTGTCTCCCTATAGTCTCTTTATGAGGTCCCGGA 840  
Db 781 GCGGCTCCAGCAGCGTTACTGGGGTGTCTCCCTATAGTCTCTTTATGAGGTCCCGGA 840  
Qy 841 GCAGGAGTGTGACAGTGTATATCCCGCCAGCAGTGGGGCCATCATCGCAAGAA 900  
Db 841 GCAGGAGTGTGACAGTGTATATCCCGCCAGCAGTGGGGCCATCATCGCAAGAA 900  
Qy 901 GGGGAGACATCAAAAGCTCTCCCGTTTTCGAGCGCTCCATCAAGATTCACCAACC 960  
Db 901 GGGGAGACATCAAAAGCTCTCCCGTTTTCGAGCGCTCCATCAAGATTCACCAACC 960  
Qy 961 CGAACACCTGACTCCAAAGTTCGTATGTTATCATCTGAGCGCGGAGGCGCAATT 1020  
Db 961 CGAACACCTGACTCCAAAGTTCGTATGTTATCATCTGAGCGCGGAGGCGCAATT 1020  
Qy 1021 CAAGGCTCAGGGAAGATCTATGGCAAACTCAAGGAGGAGAACTTTTGGTCCCAAGGA 1080  
Db 1021 CAAGGCTCAGGGAAGATCTATGGCAAACTCAAGGAGGAGAACTTTTGGTCCCAAGGA 1080  
Qy 1081 GGAAGTGAAGTGGAGACCCACATACGTGTGCGAGCATCAGCAGTGGCCGGGTCAATTGG 1140  
Db 1081 GGAAGTGAAGTGGAGACCCACATACGTGTGCGAGCATCAGCAGTGGCCGGGTCAATTGG 1140  
Qy 1141 CAAGGTGGAAGAACGGTGAACAGTTTCAGAAATTTGACGCGCAGTGAAGTGTAGTACC 1200  
Db 1141 CAAGGTGGAAGAACGGTGAACAGTTTCAGAAATTTGACGCGCAGTGAAGTGTAGTACC 1200  
Qy 1201 AAGAGACAGACCCCTGATGAGAACGACAGGTCACTGTAAGAAATCATCGGACATTTCTA 1260  
Db 1201 AAGAGACAGACCCCTGATGAGAACGACAGGTCACTGTAAGAAATCATCGGACATTTCTA 1260  
Qy 1261 TGCCAGTCAGATGGCTCAACGAGATCCGAGACATCTGTCGCCAGGTTAAGCAGCAGCA 1320  
Db 1261 TGCCAGTCAGATGGCTCAACGAGATCCGAGACATCTGTCGCCAGGTTAAGCAGCAGCA 1320  
Qy 1321 TCAGAGGAGCAGAGTAAACCGAGGCCAGGAGGAAGTGAACAGCCCTCCCTGTCC 1380  
Db 1321 TCAGAGGAGCAGAGTAAACCGAGGCCAGGAGGAAGTGAACAGCCCTCCCTGTCC 1380

## RESULT 3

US-09-061-709-7

; Sequence 7, Application US/09061709B

; Patent No. 6297364

; GENERAL INFORMATION:

; APPLICANT: Chen, Yao-Tseng

; APPLICANT: Gure, Ali

; APPLICANT: Tsang, Solam

; APPLICANT: Stockert, Elisabeth

; APPLICANT: Jager, Elke

; APPLICANT: Knuth, Alexander

; APPLICANT: Old, Lloyd J.

; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated

; TITLE OF INVENTION: Antigens, The Antigens Per Se, And Uses Thereof

; FILE REFERENCE: LUD 5538

; CURRENT APPLICATION NUMBER: US/09/061,709B

; CURRENT FILING DATE: 1998-04-17

; NUMBER OF SEQ ID NOS: 8

; SEQ ID NO 7

; LENGTH: 1946

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

US-09-061-709-7

Query Match 97.7%; Score 1669.2; DB 3; Length 1946;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 1676; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 30 CCGGGGAGCCATCATGAAGCTGAATGGCCACAGTTGGAGAACCATGCGCTGAAGGTCT 89

Db 268 CCAGGCAAGCCATCATGAAGCTGAATGGCCACAGTTGGAGAACCATGCGCTGAAGGTCT 327

Qy 90 CCTACATCCCGATGACAGATAGCACAGGACCTGAGAAATGGCGCGGCGGCTTTG 149

Db 328 CCTACATCCCGATGACAGATAGCACAGGACCTGAGAAATGGCGCGGCGGCTTTG 387

Qy 150 GCTCTCGGGGTGAGCGCCCGCCAGGGCTCACCTGTGGCAGCGGGGGCCCGCAGCAAGCAGC 209

Db 388 GCTCTCGGGGTGAGCGCCCGCCAGGGCTCACCTGTGGCAGCGGGGGCCCGCAGCAAGCAGC 447

Qy 210 AGCAAGTGGACATCCCGCTTGGCTCTGGTGGCCACCCAGATGTGGGTGCCATTATTG 269

Db 448 AGCAAGTGGACATCCCGCTTGGCTCTGGTGGCCACCCAGATGTGGGTGCCATTATTG 507

Qy 270 GCAAGGAGGGGGCCACCTCCGCAACATCAAAAACAGACCCAGTCCAAAGATAGACGTGC 329



QY 330 ATAGGAAGAGAACGACAGGTGAGTGAAGAAAGCCATCAGTGTGCACTTCCACCCCTGAGG 389  
Db |||||  
QY 568 ATAGGAAGAGAACGACAGGTGAGTGAAGAAAGCCATCAGTGTGCACTTCCACCCCTGAGG 627  
Db |||||  
QY 390 GCTGCTCTCCGCTGTTGAAGATGATCTTGGAGATATGCAATAAGAGGCTAAGACACCA 449  
Db |||||  
QY 628 GCTGCTCTCCGCTGTTGAAGATGATCTTGGAGATATGCAATAAGAGGCTAAGACACCA 687  
Db |||||  
QY 450 AAACGGCTGAGAGGTTCCCTGAGATCTCTGGCCCAATAAATCTTTGTAGGGCTGTCA 509  
Db |||||  
QY 688 AAACGGCTGAGAGGTTCCCTGAGATCTCTGGCCCAATAAATCTTTGTAGGGCTGTCA 747  
Db |||||  
QY 510 TTGGCAAGAGAGGACGGAACCTGAAGAGAGGTAGAGCAAGATACCGAGACAAATAACCA 569  
Db |||||  
QY 748 TTGGCAAGAGAGGACGGAACCTGAAGAGAGGTAGAGCAAGATACCGAGACAAATAACCA 807  
Db |||||  
QY 570 TCTCTCTGTTGCAAGACCTTACCTTTTAAACCTCTGAGAGGACCATCATCTGTGAAGGGG 629  
Db |||||  
QY 808 TCTCTCTGTTGCAAGACCTTACCTTTTAAACCTCTGAGAGGACCATCATCTGTGAAGGGG 867  
Db |||||  
QY 630 CCATCGAATTTGTGAGGGCCGAGCAGGAATAATGAAGAAAGTTTGGGAGGCTTATG 689  
Db |||||  
QY 868 CCATCGAATTTGTGAGGGCCGAGCAGGAATAATGAAGAAAGTTTGGGAGGCTTATG 927  
Db |||||  
QY 690 AGAATGATGGCTGCGATGAGCTCTACCTGATCCCTGGCTGACCTGGCTGCTGTAG 749  
Db |||||  
QY 928 AGAATGATGGCTGCGATGAGCTCTACCTGATCCCTGGCTGACCTGGCTGCTGTAG 987  
Db |||||  
QY 750 GTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCCGCTCCAGAGAGGCTTACTGGGGCTG 809  
Db |||||  
QY 988 GTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCCGCTCCAGAGAGGCTTACTGGGGCTG 1047  
Db |||||  
QY 810 CTCCTATAGCTCTTTATGAGGCTCCCGAGGAGGAGATGGTGCAGGTGTTTATCCCGG 869  
Db |||||  
QY 1048 CTCCTATAGCTCTTTATGAGGCTCCCGAGGAGGAGATGGTGCAGGTGTTTATCCCGG 1107  
Db |||||  
QY 870 CCCAGGAGTGGGGCCCATCATCGCAAGAGAGGGGACACATCAAAACAGCTCTCCCGG 929  
Db |||||  
QY 1108 CCCAGGAGTGGGGCCCATCATCGCAAGAGAGGGGACACATCAAAACAGCTCTCCCGG 1167  
Db |||||  
QY 930 TTGCCAGGCTCCATCAAGATTTGCACACCCGAAACCTGACTCCAAAGTTTGGTATGG 989  
Db |||||  
QY 1168 TTGCCAGGCTCCATCAAGATTTGCACACCCGAAACCTGACTCCAAAGTTTGGTATGG 1227  
Db |||||  
QY 990 TTATCATCACTGGACCGGAGGCGCAATTCAGGCTCAGGAGAGATCTATGGCAAC 1049  
Db |||||  
QY 1228 TTATCATCACTGGACCGGAGGCGCAATTCAGGCTCAGGAGAGATCTATGGCAAC 1287  
Db |||||  
QY 1050 TCAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAGTGAAGCTGGAGACCCACATAGCTG 1109  
Db |||||  
QY 1288 TCAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAGTGAAGCTGGAGACCCACATAGCTG 1347  
Db |||||  
QY 1110 TGCAGCATCAGCAGCTGGCGGGTCATTTGGCAAAAGGTGGAAACCGGTGAACGAGTTGC 1169  
Db |||||  
QY 1348 TGCAGCATCAGCAGCTGGCGGGTCATTTGGCAAAAGGTGGAAACCGGTGAACGAGTTGC 1407  
Db |||||  
QY 1170 AGAATTTGACGGCAGCTCAGGTGGTGGTACCAAGAGACGACCCCTGATGAGACGAC 1229  
Db |||||  
QY 1408 AGAATTTGACGGCAGCTCAGGTGGTGGTACCAAGAGACGACCCCTGATGAGACGAC 1467  
Db |||||  
QY 1230 AGGTCACTCGTGAATTCATCGGACATTTCTATGCGAGTCAAGTGGCTCAACCGGAGTCC 1289  
Db |||||  
QY 1468 AGGTCACTCGTGAATTCATCGGACATTTCTATGCGAGTCAAGTGGCTCAACCGGAGTCC 1527  
Db |||||  
QY 1290 GAGACATCTGCGCCAGGTTAAGCAGCAGCATCAGAGGGACAGTAAACCGCCGAG 1349  
Db |||||  
QY 1528 GAGACATCTGCGCCAGGTTAAGCAGCAGCATCAGAGGGACAGTAAACCGCCGAG 1587  
Db |||||  
QY 1350 CACGGAGGAAGTGAACGACCCCTCTGCTGCTTNGAGTCCAGGACCAACACCGGCGAGAA 1409  
Db |||||  
QY 1588 CACGGAGGAAGTGAACGACCCCTCTGCTGCTTNGAGTCCAGGACCAACACCGGCGAGAA 1647  
Db |||||

QY 1410 ATCAGAGTGTGCTCTCCCGGAGCGCTGAGAAATGAGTGGGAATCCGGGACACNCTGGGC 1469  
Db |||||  
QY 1648 ATCAGAGTGTGCTCTCCCGGAGCGCTGAGAAATGAGTGGGAATCCGGGACACNCTGGGC 1707  
Db |||||  
QY 1470 CGGGCTGTAGATCAGGCTTTCACCTTGAATGAGAAAGATGTTCCAGTGAAGAACCTCGA 1529  
Db |||||  
QY 1708 CGGGCTGTAGATCAGGCTTTCACCTTGAATGAGAAAGATGTTCCAGTGAAGAACCTCGA 1767  
Db |||||  
QY 1530 TCTNTCAGCCCCAAAACACCCCAATTCGCCCAACACTGTNTGCCCTCGGGGTGTGAG 1589  
Db |||||  
QY 1768 TCTNTCAGCCCCAAAACACCCCAATTCGCCCAACACTGTNTGCCCTCGGGGTGTGAG 1827  
Db |||||  
QY 1590 AAATTTAGGGCAAGCAGCTTTAAACGTGGATGTTTAAAGAGCTCTCCAGGCCCCAC 1649  
Db |||||  
QY 1828 AAATTTAGGGCAAGCAGCTTTAAACGTGGATGTTTAAAGAGCTCTCCAGGCCCCAC 1887  
Db |||||  
QY 1650 CAAGAGGGTGGATCAGACCTCAGTGGGAAAGAAATAAAATTTCCCTTCAGGTTTAAAA 1708  
Db |||||  
QY 1888 CAAGAGGGTGGATCAGACCTCAGTGGGAAAGAAATAAAATTTCCCTTCAGGTTTAAAA 1946  
Db |||||

RESULT 5  
US-09-261-855-1  
; Sequence 1, Application US/09261855A  
; Patent No. 6255055  
; GENERAL INFORMATION:  
; APPLICANT: Ross, Jeffrey  
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN  
; FILE REFERENCE: 960296.95131  
; CURRENT APPLICATION NUMBER: US/09/261,855A  
; CURRENT FILING DATE: 1999-03-03  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 2224  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-261-855-1

Query Match 68.2%; Score 1165; DB 3; Length 2224;  
Best Local Similarity 82.8%; Pred. No. 0;  
Matches 1411; Conservative 0; Mismatches 260; Indels 33; Gaps 6;

QY 30 CCCGGGAGGCATCATGAAGCTGAATGGCCACACAGTTGGAGAACCATGCCCTGAAGCTCT 89  
Db |||||  
QY 525 CCAGGCAAGCTATCATGAAGCTAAATGGCCATCACTGGAGAACCATGCCCTGAAGCTCT 584  
Db |||||  
QY 90 CCTACATCCCGATGAGCAGATAGCACAGGACCTGAGAAATGGCGCCGAGGGGCTTTG 149  
Db |||||  
QY 585 CCTACATACCTGATGAGCAGATAACAAAGTCTCTGAGAATGGCGCTCGTGGAGGCTTTG 644  
Db |||||  
QY 150 GCTCTGGGGTCAGCCCCCGCCAGGGCTCAGCTGTGCGAGGGGGGGCCCGCAAGCAGC 209  
Db |||||  
QY 645 GGTCTCGGGGCCAGCCCCCGCAAGGGTTCGCGCAGCAGGGGGCTCCAGCCCAAGCAGC 704  
Db |||||  
QY 210 AGCAAGTGAACATCCCGCTTCCGCTCTCGGTGCCACCCAGATGATGTGGGTGCCATTATTG 269  
Db |||||  
QY 705 AGCAGTGGACATCCCTCTCCGGCTCTCGGTGCTACGAGTATGTAGGGCTATCATTTG 764  
Db |||||  
QY 270 GCAAGAGGGGGCCACATCCCGCAACATCAAAAAACAGACCCAGTCCAAAGATAGACGTGC 329  
Db |||||  
QY 765 GCAAGAGGGGTGCCACATCCGAAACATCAAAAAACAGACCCAGTCCAAAGATAGACGTGC 824  
Db |||||  
QY 330 ATAGGAAGAGAGACGAGGTGACGCTGAGTGAAGAAAGCCATCAGTGTGCACTCCACCCCTGAGG 389  
Db |||||  
QY 825 ATAGGAAGAGAGATCGGGGCTCGGAGAGAGCCATCAGCTGTCATTCAACCCCTGAAG 884  
Db |||||  
QY 390 GCTGCTCTCCGCTGTTAAGATGATCTTGGAGATTTATGATAAAGAGGTAAAGACACCA 449  
Db |||||  
QY 885 GCTGCTCTCCGCTGCAAGATGATCTTGGAGATTTATGCAAGAGGAGGAAAGACACCA 944  
Db |||||  
QY 450 AAACGGCTGACAGGTTCCCTGAAAGATCCTGCCCATTAATACTTTGTAGGCGCTCTCA 509  
Db |||||

Db 945 AAACGGCAGATGAAGTTCCCTTGAAGATCCCTGGCTCATAAACATTCGTGGGGGACTCA 1004  
Qy 510 TTGGCAAGGAAGGAAGGAACCTTGAAGAAGGTAGAGCAAGATACCGAGACAAATATCACCA 569  
Db 1005 TTGGCAAGGAAGGAAGGAACCTTGAAGAAGGTAGAGCAAGATACCGAGACAAATATCACCA 1064  
Qy 570 TCTCTCTGTTGCAAGACCTTACCTTTTACACCTTGAAGAGACATCATCTGTGAAGGGGG 629  
Db 1065 TCTCATCGTCCAGGACCTCAGCTCTATAACCTTGAAGAGACATCATCTGTGAAGGGGG 1124  
Qy 630 CCATCGAGAAATTTGTCAGGGGCGGAGCAGGAAATTAATGAAGAAATTCGGGAGGCTATG 689  
Db 1125 CCATTGAGAACTGTTGCGAGGCGGAGCAGGAGATCATGAAGAAATTCGAGAGGCTTACG 1184  
Qy 690 AGAATGATGTGGCTGCCATGAGC-----TCTACCTGATCCCTGGCTGAACCTGGCTG 743  
Db 1185 AGAAGCAGGTGGCGGCCATGAGCTTGAGTCCAGTCCACCTCATCCCTGGGCTTAACTGGCTG 1244  
Qy 744 CTGTAGGTCTTTTCCAGCTTCATCCAGGAGTCCCGCCGCTCCAGCAGCTTACTG 803  
Db 1245 CTGTAGGTCTTTTCCAGCTTCATCCAGGAGTCCCGCCGCTCCAGCAGCTTACTG 1304  
Qy 804 GGGCTGCTCCCTATAGCTCCCTTTATGSCAGGCTCCCGAGCAGGAGATGGTGAGGTGTTA 863  
Db 1305 GGGCTGCTCCCTATAGCTCCCTCATGAGGCTCCCGAGCAGGAGATGGTCAAGTGTCA 1364  
Qy 864 TCCCGCCGAGCAGTGGCGGCATCATCGGAAGAGGGGAGCAGACATCAACAGCTCT 923  
Db 1365 TCCCGCCGAGCAGTGGCGGCATCATCGGAAGAGGGGAGCAGACATCAACAGCTCT 1424  
Qy 924 CCGGTTTGGCAGCGCTCCATCAAGATTGACACCGCAACCTGACTCCAAAGTTC 983  
Db 1425 CCGGTTTGGCAGCGCTCCATCAGATTGTCACAGAAACCTGACTCCAAAGTTC 1484  
Qy 984 GTATGGTTATCATCTAGCGCGCAGAGGCCCAATTCAGGCTCAGGGAAGATCATG 1043  
Db 1485 GAAATGGTCTCATCTAGCGCGCCAGAGGCTCAGTTCAGGCTCAGGGAAGATTTATG 1544  
Qy 1044 GCAAACTCAAGGAGAGAACTCTTTGTCCTCAAGGAGGAGTCAAGCTGAGACCCACA 1103  
Db 1545 GCAAACTCAAGGAGAGAACTCTTTGTCCTCAAGGAGGAGTCAAGCTGAGACCCACA 1604  
Qy 1104 TACGTGTCCAGCATCAGCAGTGGCGGGTCAATGGCAAGGTGGAACCGTGAACG 1163  
Db 1605 TACGGTTTCCGGTTTACGAGCGCGCGTCAATGGCAAGGTGGAACCGTGAACG 1664  
Qy 1164 AGTTGCAAGTTTGAAGGAGTGGGTGAGTACCAAGAGACAGACCCCTGATGAGA 1223  
Db 1665 AGCTCGAAGCTTGACTCAGTGGGTGAGTGGCAAGAGACAGACCCCGATGAGA 1724  
Qy 1224 ACGACAGGTCATCGTGAATTCATCGGACATTTCTATGCCAGTCAAGTGGCTCAACGGA 1283  
Db 1725 ACGACCAAGTCATTTGAAGATCATCGGACATTTCTATGCCAGCAGATGGCTCAGCGGA 1784  
Qy 1284 AGATCCGAGACATCTGGCCAGGTTAAGCAGAGCATCAGAGGGAAGATGACAGG 1343  
Db 1785 AGATCCGAGACATCTGGCTCAAGTTAAGCAACAGCACCAGAGGGAAGATGACAGG 1844  
Qy 1344 CCCAGGACGAGGAGAGTGA--CCAGCCCTCCCTGTCCTTNGAGTCCAGGACCAACAG 1402  
Db 1845 CCCAGGACGAGGAGAGTGA--CCAGCCCTCCCTGTCCTTNGAGTCCAGGACCAACAG 1904  
Qy 1403 GCGAGAA-----ATCGAGAGTGTGCTCTCCCGGAGGCTGAG 1441  
Db 1905 AGGAACACAGAACTGAGGGGGGGTGGAGGGCGGTGTGTTTTTCCAGCAGGCTGAG 1964  
Qy 1442 AATGAGTGGGAATCCGGGACATNTGGGCGGGCTGTAGATCAGGTTTGGCCACTTGATTG 1501  
Db 1965 AATGAGTGGGAATCAGGG--CATTTGGGCTGGCTGGAGATCAGGTTTGGCACTGATG 2023  
Qy 1502 AGAAGATGTTCCAGTGAAGAACCTGATCTNTCAGCCCAACCCCAACCAATTGGCC 1561

Db 2024 AGAACAAATTTCCAGTGAAGAAATCTGTATCTCTCGCCCCCAA--TTGAGCCAGCTGGCCA 2081  
Qy 1562 CAACACTGNTGCCCCCTCGGGGTGTCAAAAATNTAGCCCAAGGCACTTTTAAACGTGGA 1621  
Db 2082 CAGCCCAACCCCTTGAATATCACCATTGCAATCATAGCTTTGGTTGCTTTTAAACGTGGA 2141  
Qy 1622 TTGTTTAAAGAGCTTCCAGGCCCCCACCAGAGGGTGGATCACACTTCAGTGGGAAGAA 1681  
Db 2142 TTGTCT--TGAAGTTCTCCAGCTCCATGGAAGGATGGGTCCAGATCCCAGTGGGAAGAG 2199  
Qy 1682 AAATAAAATTTCTTCAGGTTTTA 1705  
Db 2200 AAATAAAATTTCTTCAGGTTTTA 2223

RESULT 6  
US-09-643-597-347  
; Sequence 347, Application US/09643597  
; Patent No. 6426072  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: McNeill, Patricia D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.455C11  
; CURRENT APPLICATION NUMBER: US/09/643,597  
; CURRENT FILING DATE: 2000-08-21  
; NUMBER OF SEQ ID NOS: 369  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 347  
; LENGTH: 1740  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-643-597-347

Query Match 34.1%; Score 582.2; DB 4; Length 1740;  
Best Local Similarity 66.2%; Pred. No. 2.8e-156;  
Matches 858; Conservative 0; Mismatches 433; Indels 6; Gaps 1;

Qy 37 AGCCATCATGAAGCTGAATGGCCACAGTGGAGAACCATGCCCTGAAGGCTCTCTACAT 96  
Db 402 AGCACTAGACAACTGAATGGATTTTCAGTTAGAGAAATTTACCTTGAAGTAGGCTATAT 461  
Qy 97 CCCCAGATGAGCAGATAGCACAGGACCTGAGAAATGGCGCGGAGGGGCTTTGGCTCTCG 156  
Db 462 CCTGATGAAACGGCGCCAGCAAAACCCCTTGAGCAGAGCCCGAGTCCGCCGGGGCT 521  
Qy 157 GGGTCAGCCCCCGCAGGGGCTCCTGTGGCAGCGGGGCCCCAGCCAGCAGCAGCAAGT 216  
Db 522 TGGGAGAGGGGGCTCCTCAAGGAGGGGTCTCCAGGATCCGCTATCAAGAGCAAAACCATG 581  
Qy 217 GACATCCCTTCCTGGCTCTGGTCCCAAGGAGTATGGTGGCTATTTTGGCAAGGA 276  
Db 582 TGATTGGCTCTGGCCCTGCTGGTTCCCAAGGAGTATTTGTTGAGGCATCATAGGAAAGA 641  
Qy 277 GGGGGCCACCATCCCAACATCAAAAACAGACCCAGTCCCAAGATAGAGTAGGATAGAA 336  
Db 642 AGGTGCCACCATTCGGAACATCAACCAACAGACCCAGCTTAAATCGATGTCCACCGTAA 701  
Qy 337 GAGAACGAGGTCAGCTGAGGAAAGCCATCAGTGTGACCTCCACCCCTGAGGGTGTCTC 396  
Db 702 AGAAATCGGGGGCTGTGAGAGTCAATTAATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 761  
Qy 397 CTCCGCTTGAAGATGATCTTTGAGATTATGATTAAGAGGCTAAGGACACCAAAACCGC 456



Db 762 TCGGCTGTAGTCTATTCTGGAGATTATGCTAAGAGCTCAAGATATAAAATTCAC 821  
QY 457 TCACGAGTTCCCTGGAAGATCCTGCCCATATAAATCTTGTAGGCGTCTCATTTGGCAA 516  
Db 822 AGAAGAGATCCCTTGAAGATTTAGCTCATAATAAATCTTGTGGACGCTTATTGGTAA 881  
QY 517 GGAAGAGCGACCTTGAAGAGTGTAGCAGATACCGAGACACAAAATCAACATCTCTC 576  
Db 882 AGAAGAGAAATCTTAAAAAATTTGAGCAAGACACACATATAAATCAGATATCTCC 941  
QY 577 GTTGAAGACCTTACCCCTTTACAACCTTGAGAGGACCATCACTGTGAAGGGGGCCATCGA 636  
Db 942 ATTGCAGGAATTCAGCTGTATATAATCCAGAACGCACTATTACAGTTAAAGCAATGTTGA 1001  
QY 637 GAATTTGTCAGGGCGGACGAGAAATATGAGAAAGTTGCGGAGGCTTATGAGATGA 696  
Db 1002 GACATGTGCCAAGCTGAGGAGAGATCATGAAGAAATATCAGGAGTCTTATGAAAAATGA 1061  
QY 697 TGTGGCTGCATGA-----GCTCTCACCTGATCCTGGCCTGAACCTGGCTGCTGTAGG 750  
Db 1062 TATTGCTTCTATGATCTTCAGACACATTTAATTCCTGGATTAATCTGAACGCTTGGG 1121  
QY 751 TCTTTTCCAGGTTCAATCAGCGCATGTCCTGGCTCCCGGCTCCAGCAGGTTACTTGGGGTGC 810  
Db 1122 TCTGTTCCACCACTTCAGGGATGCCACCTCCACCTCAGGGCCCCCTTCAGCCATGAC 1181  
QY 811 TCCCTATAGTCTTTATGACGCTCCGAGCAGGATGTCAGAGTGTATATCCCGC 870  
Db 1182 TCTCTCCCTACCGCAGTTTGAGCAATCAGAAACGAGACTGTTTATCTGTTATCCACG 1241  
QY 871 CCAGGACGTGGCGCATCATCGCAAGAGGGGAGCACATCAACACAGCTCTCCGGTT 930  
Db 1242 TCTATCAGTCGGTGCCATCATCGCAAGCAGGGCCAGCACATCAAGCAGCTTTCGCTT 1301  
QY 931 TCCAGCGCTCCATCAGATTGACACCCGACCCGAAACACCTGACTCCAAAGTTCTGATGGT 990  
Db 1302 TCGTGGAGCTTCAATTAAGATTGCTCCAGCGGAGCCAGATGCTTAAAGTGAGGATGGT 1361  
QY 991 TATCATCACTGACCGCCAGAGGCCCAATTTCAAGGCTCAGGGAAGAACTTATGGCAACT 1050  
Db 1362 GATTATCACTGACCAACAGAGCTCAGTTCAAGGCTCAGGGAAGAAATTTATGNAAT 1421  
QY 1051 CAAGGAGGAGACTTCTTTGTCCTCAAGAGGAAAGTGAAGTGGAGACCCATACAGTGT 1110  
Db 1422 TAAAGAGAAACTTTGTTAGTCTTAAAGAGAGGTGAACTTTGAAGCTCATATCAGAGT 1481  
QY 1111 GCAGCATCAGCAGCTGGCGGCTCATTTGGCAAGGTGGAAGACGGTGAACGAGTTGCA 1170  
Db 1482 GCCATCTTGTCTGGCAGAGTTATTGGAAAGAGGCAAAACGGTGAATGAATTTCA 1541  
QY 1171 GAATTTGACGGCAGCTGAGGTGTATGATACCAAGAGACGAGACCCCTGATGAGAACGCA 1230  
Db 1542 GAATTTGTCAAGTGAGAGTGTGTTGCTCCCTGTCAGCAGACACCTGATGAGATGACA 1601  
QY 1231 GGTATCGTGAATATCAGGACATTTCTATCCAGTCAAGTGGCTCAACGGAAGATCG 1290  
Db 1602 AGTGGTTGCAAAATAACTGGTCACTTTATCTGCTTGCAGAGTTGCCAGAGAAATTTCA 1661  
QY 1291 AGACATCTCTGGCCAGGTTAAGCAGCAGCATCAGAAAG 1327  
Db 1662 GGAATTTCTGACTCAGGTTAAGCAGACCAACACAG 1698

RESULT 7

US-09-542-615A-347  
; Sequence 347, Application US/09542615A  
; Patent No. 6518256  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: Hosken, Nancy A.  
; APPLICANT: Fanger, Gary R.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C8  
; CURRENT APPLICATION NUMBER: US/09/542,615A  
; CURRENT FILING DATE: 2000-04-14  
; NUMBER OF SEQ ID NOS: 350  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 347  
; LENGTH: 1740  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-542-615A-347

Query Match 34.1%; Score 582.2; DB 4; Length 1740;  
Best Local Similarity 66.2%; Pred. No. 2.8e-156;  
Matches 859; Conservative 0; Mismatches 433; Indels 6; Gaps 1;

QY 37 AGCCATCATCAAGCTGAATGGCCACACAGTTGGAGAACCATGCTCAAGGCTCTCCATACAT 96  
Db 402 AGACTAGACAACTGATGATTTAGTTAGAGATTTCACTTGAAGTAGCCTATAT 461  
QY 97 CCCGATGACGATAGCACAGGAGCCTGAGAAATGGCGCCGAGGGGGCTTTGGCTCTCG 156  
Db 462 CCCTGATGAACCGCGCCCGCAGCAAAACCCCTTGCAGCAGCCCGAGGTGCGCGGGCT 521  
QY 157 GGGTCAGCCCCCGCAGGGCTCACTGTGGCAGCGGGGGCCCCAGCCAGCAGCAGCAAGT 216  
Db 522 TGGGACGAGGGGCTCTCAAGGACGGGCTCTCCAGGATCCGTATCCAGCAGAAACCATG 581  
QY 217 GGCATCCCCCTTCGGCTCCTGTGTGCCACCCAGTATGTGGTGCCATTTATGGCAAGGA 276  
Db 582 TGAATTTGCTCTGCGCTGCTGCTGCCACCCCAATTTGTTGGAGCCATCATAGNAAAGA 641  
QY 277 GGGGCCACCATCGGCAACATCAAAAACAGACCCAGTCCAGATAGACATAGCATAGGAA 336  
Db 642 AGGTGCCACCATTCGGAAACATCAAAAACAGACCCAGTCTAAAATCGATCTCCACCGTAA 701  
QY 337 GGAAGACGAGGTGCAGCTGAAAAGCCATCAAGTGTGCTCACTCCACCCCTGAGGCTGCTC 396  
Db 702 AGAAATGCGGGGCTGCTGAGAGTGTACTTCTCTACTCTCTGAGGACCTC 761  
QY 397 CTCGGTTGTAAGATGATCTTGGAGATTTATGCATAAAGAGGTAAAGACACCAAAACGGC 456  
Db 762 TGGGGTTGTAAGTCTATTCTGGAGATTTATGCATAAAGAGGTAAAGACATATAAAATTCAC 821  
QY 457 TGACGAGTTCCCTGGAAGATCCTGGCCATATAAATTTGTAGGGGCTCTCATTTGGCAA 516  
Db 822 AGAAGATCCCTTGAAGATTTAGCTATAATAAATTTGTTGGAGCTTTATTTGGTAA 881  
QY 517 GGAAGACGGAACCTGAAAGAGGTAGAGCAAGATACCGAGACAAAATCAACATCTCTCCTC 576  
Db 882 AGAAGGAAGAAATCTTAAAAAATTTGAGCAAGACACAGACACTAAAATCAGATATCTCC 941  
QY 577 GTTCAAGACCTTACCCCTTTACAACTGAGAGGACCATCACTGTAGGGGGCCATCGA 636  
Db 942 ATTGCAGGAATGACGCTGTATAATCCAGAACCACTATTACAGTTAAAGGCAATGTTGA 1001  
QY 637 GAATTTGTCAGGGCCGAGCAGGAATAATGAAGAAAGTTGCGGAGGCTTATGAGATGA 696  
Db 1002 GACATGTGCCAAGCTGAGGAGGATCATGAGAAATCAGGAGTCTTATGAAATGA 1061  
QY 697 TGTGGCTGCCATGA-----GCTCTCACCTGATCCTTGGCCTGAACTGGCTGCTGTAGG 750  
Db 1062 TATTGCTTCTATGAATCTTCAAGCACAATTTAATTCCTGGATTAATCTGAACGCTTGGG 1121  
QY 751 TCTTTTCCAGCTTCAATCCAGCGCAGTCCCGCGCTCCAGCAGAGGTTACTTGGGGTGC 810  
Db 1122 TCTGTTCCCACTTCAGGATGCACTCCCACTCCAGCTCAGGGCCCCCTTCAGCCATGAC 1181  
QY 811 TCCCTATAGTCTCTTATGACGCTCCGAGCAGGATGTCAGAGTGTATATCCCGC 870

Db 1182 TCCTCCCTACCGCGAGTTTGAAGCAATCAGAAACGGAGAGCTGTTTCATCTGTTTATCCCAGC 1241  
QY 871 CCAGGCGAGTGGCGGCATCATCGCAAGAGGGGAGACATCAAAACAGCTCTCCCGGT 930  
Db 1242 TCATCAGTCGGTGCATCATCGCAAGAGGGGAGACATCAAGCAGCTTTCGCTT 1301  
QY 931 TCACAGCGCTCCATCAAGATGCAACCCGAAACACCTGACTCCAAAGTTGCTATGGT 990  
Db 1302 TGTGAGGCTTCAATTAAGATTGCTCCAGCGAAGCACAAGATGCTAAAGTGAGGATGGT 1361  
QY 991 TATCATCTAGCAGCGCCAGAGGCCCAATTCAGAGCTCAGGGAAGATCTATGCAACT 1050  
Db 1362 GAATATCACTGGACCAACAGAGCTCAGTTCAGAGCTCAGGGAAGATTTATGGAAGAT 1421  
QY 1051 CAAGGAGGAGAACTTCTTTGGTCCCAAGAGGAAGTGAAGCTGGAGACCCACATACGTTG 1110  
Db 1422 TAAAGAGAGAACTTTGTAGTCTTAAAGAGAGGTTGAACCTTGAAGCTCATATCAGAGT 1481  
QY 1111 GCCAGATCAGAGCTGGCGCGGTCAATTGGCAAGGTGGAAGCGTGAAGTGGCA 1170  
Db 1482 GCCATCTTGTCTGCTGGCAGAGTTATTGGAAAGGAGGCAAAACGGTGAATGAATCA 1541  
QY 1171 GAATTTGACGGCAGCTGAGTGGTAGTACCAAGAGACAGACCCCTGATGAGAAGCA 1230  
Db 1542 GAATTTGCAAGTGCAGAGTTGTTGCTCGTGACCAAGACCTGATGAGATGACCA 1601  
QY 1231 GGTATCGTGAAGATCATCGGACATTTCTATGCCAGTCAAGTGGCTCAACGGAAGATCCG 1290  
Db 1602 AGTGGTTGTCAAAATAAAGTGGTCACTTCTATGCTTGCAGGTTGCCAGAGGAGGCA 1661  
QY 1291 AGACATCTGCGCCAGGTTTACGACGACATCAGAG 1327  
Db 1662 GGAATTTGACTCAGTAAAGCAGCAGCAACACAG 1698

## RESULT 8

US-09-606-421B-347  
; Sequence 347, Application US/09606421B  
; Patent No 6531315  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy  
; APPLICANT: Panger, Gary R.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 2101.21.455C9  
; CURRENT APPLICATION NUMBER: US/09/606,421B  
; NUMBER OF SEQ ID NOS: 358  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 347  
; LENGTH: 1740  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-606-421B-347

Query Match 34.1%; Score 582.2; DB 4; Length 1740;  
Best Local Similarity 66.2%; Pred. No. 2.8e-156;  
Matches 858; Conservative 0; Mismatches 433; Indels 6; Gaps 1;  
QY 37 ACCCATCATGAGCTGAATGGCCACCACTGGAGAACCATGCTCCCTGAAGTCTCTCAT 96  
Db 402 AGCACTAGCAAACTGAATGATTTCAAGTATGAGAAATTCACCTTGAAGTAGCCTAT 461  
QY 97 CCCCATGAGCAGATGACACAGGACCTGAGAATGGCGCCGAGGGGCTTTGGCTCTCG 156  
Db 462 CCTGATGAACGGCGCCAGCAAAACCCCTTGCAGCAGCCCGAGGCTCGCGGGGCT 521

QY 157 GGCTCAGCCCCGCGACGGGCTCACCTGTGCGAGCGGGGGCCCCAGCAACGACGACGCACT 216  
Db 522 TGGGAGAGGGGCTCTCAAGGAGGGGCTCCAGGATCCGTATCCAGCAGAAACCATG 581  
QY 217 GGCATATCCCCCTTCGCTCTCGTGGTCCCAACCAAGTATGTTGGTGGCAATTTATGGCAAGA 276  
Db 582 TGATTTGCTCTCGGCTCTGCTGTTCCCAACCAATTTGTTGGAGCATCATAGGAAAGA 641  
QY 277 GGGGGCCACCTCCGCAACATCACAAAACAGACCCAGTCCAGATAGAGATAGAGTGCATAGGA 336  
Db 642 AGTGCCACCATTCGGAACATCACAAACAGACCCAGTCTAAATTCGATGTCCACCGTAA 701  
QY 337 GGAAGACGAGGTGACGTGAAAGGCCATCATGTCGACTCCACCCCTGAGGGGTGCTC 396  
Db 702 AGAAATGCGGGGCTGCTGAGAGTCGATTACTATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 761  
QY 397 CTCGGCTTGAAGATCTTGGAGTTATGATTAAGAGGCTAAGGACACCAAAACGGC 456  
Db 762 TGGGGCTTGAAGTCTATTCTGGAGATTATGATTAAGGAAGCTCAAGATATAAAATTCAC 821  
QY 457 TGACGAGGTTCCCTCGAAGATCCTGCGCCATATAAATTTGTAGGGCGTCTCATTTGGCAA 516  
Db 822 AGAGAGATCCCTTGAAGATTTAGTCTCATATAAATTTGTTGGACGCTCTTATTTGTTAA 881  
QY 517 GGAAGACGGAACCTGAAGAGGTAGCAAGATACCGAGACAAATAACCATCTCTCTC 576  
Db 882 AGAAGAGAAATCTTAAATAAATTTGAGCAAGACACAGACACTAAATCACGATATCTCC 941  
QY 577 GTTGAAGACCTTACCTTTTACAACTCGAGAGGACCATCTGTCGAGGGGGCCATCGA 636  
Db 942 ATTGAGGAATTGACGCTGATATTCAGAACGCACTTATACAGTTAAGGCAATGTTGA 1001  
QY 637 GAATTTGCGAGGGCGGACGAGCAAAATATAGAAAGTTTCGGAGGGCTATGCAATGA 696  
Db 1002 GACATGTGCCAAAGCTGAGGAGGAGATCATGAAGAAATCAGGGAGTCTTATGAATAATGA 1061  
QY 697 TGTGGCTGCATGA-----GCTCTCAGCTGATCCCTGGCTGACCTGGCTGGCTGAGG 750  
Db 1062 TATTGCTTCTATGAATCTTCAAGCAGATTAATCTCTGATTAATAATCTGAACGCTTGG 1121  
QY 751 TCTTTTCCAGGTTTATCCAGCGCAGTCCCGCGGCTCCAGCAGCGTTACTTGGGGCTGC 810  
Db 1122 TCTGTTCCCAACCACTTCAGGGATGCCACTCCACCTCAGGGCCCTTCAGCCATGAC 1181  
QY 811 TCCTATAGCTCTTATGACAGCTCCGAGCAGAGATGTTGAGGCTGTTTATCCCGC 870  
Db 1182 TCTCTCTACCGCAGTTTGAAGATTCAGAAACGGAGACTGTTCTATCTGTTTATCCGAC 1241  
QY 871 CCAGGAGTGGCGGCATCATCGCAAGAGGGGAGACATCAAAACAGCTCTCCCGGT 930  
Db 1242 TCTATCAGTCGTCGATCATCGCAAGAGGGCCAGCAGCATCAAGCAGCTTCTCGCT 1301  
QY 931 TGCAGCGCTCCATCAGATTGCAACCCGAAACACTGACTCCAAAGTTGCTATGGT 990  
Db 1302 TGTGGAGCTTCAATTAAGATTGCTCCAGCGGAAGCAGATGCTAAAGTGAGGATGGT 1361  
QY 991 TATCATCTGGACCGCGAGGCGCAATTCAGAGCTCAGGGAAGATCTATGGAAGCT 1050  
Db 1362 GAATATCACTGGACCAACAGAGGCTCAGTTCAAGGCTCAGGGAAGATTTATGGAAGAT 1421  
QY 1051 CAAGGAGGAGAACTTTTGTGTTCCCAAGAGGAAGTGAAGCTGGAGACCCCATACGCTG 1110  
Db 1422 TAAAGAGAGAACTTTGTAGTCTTAAAGAGAGGTGAACCTTGAAGCTCATATCAGAGT 1481  
QY 1111 GCCAGATCAGCAGCTGGCGGTCAATTGGCAAGGTGGAAGAGGTTGAAGTTCGA 1170  
Db 1482 GGCATCTTGTCTGCTGGCAGAGTTATTGGAAAGGAGGCAAAACGGTGAATGAATCA 1541  
QY 1171 GAATTTGAGGGGAGCTGAGTGGTAGTACCAAGAGACAGACCCCTGATGAGAAGCA 1230  
Db 1542 GAATTTGTCAAGTGCAGAGTTGTTGTCCTCTGTCAGCAGACACCTGATGAGATGACCA 1601





```

, LOCATION: (3549)
, OTHER INFORMATION: n=A,T,C or G
, NAME/KEY: unsure
, LOCATION: (3646)
, OTHER INFORMATION: n=A,T,C or G
, NAME/KEY: unsure
, LOCATION: (3940)
, OTHER INFORMATION: n=A,T,C or G
, NAME/KEY: unsure
, LOCATION: (3968)
, OTHER INFORMATION: n=A,T,C or G
, NAME/KEY: unsure
, LOCATION: (3974)
, OTHER INFORMATION: n=A,T,C or G
, NAME/KEY: unsure
, LOCATION: (4036)
, OTHER INFORMATION: n=A,T,C or G
, NAME/KEY: unsure
, LOCATION: (4056)
, OTHER INFORMATION: n=A,T,C or G
, NAME/KEY: unsure
, LOCATION: (4062)
, OTHER INFORMATION: n=A,T,C or G
, NAME/KEY: unsure
, LOCATION: (4080)
, OTHER INFORMATION: n=A,T,C or G
, NAME/KEY: unsure
, LOCATION: (4088)
, OTHER INFORMATION: n=A,T,C or G
, NAME/KEY: unsure
, LOCATION: (4115)
, OTHER INFORMATION: n=A,T,C or G
US-09-643-597-175

```

| Query Match           | 34.0%; | Score 580.6;   | DB 4;           | Length 4181;      |
|-----------------------|--------|--|-----------------|-------------------|
| Best Local Similarity | 66.1%; | Prod. No. 1.2e-155;  |                 |                   |
| Matches               | 857;   | Conservative 0;  | Mismatches 434; | Indels 6; Gaps 1; |
| Qy                    | 37     | AGCCATCATGAAGCTGAATGCCACCCAGTTGGAGAACCATGCCCTCGAAGGTCCTCCTACAT | 96              |                   |
| Db                    | 652    | AGCACTAGACAAACTGAATGATTTTCAGTTAGAGAAATTCACCTTTGAAGTAGGCTATAT   | 711             |                   |
| Qy                    | 97     | CCCCGATGAGCAGATAGCACAGGGACCTGAGAAATGGCGCGCGAGGGGGCTTTGGCTCTCG  | 156             |                   |
| Db                    | 712    | CCCTGATGAATGSCCGCCAGCAAAACCCCTTCAGCAGAGCCCGAGGTCCCGGGGGCT      | 771             |                   |
| Qy                    | 157    | GGGTGAGCCCGCCAGGGCTCACTGTGGCAGCGGGGGCCCGACGCCAAGCAGCAGCAAGT    | 216             |                   |
| Db                    | 772    | TGGCAGAGGGGCTCCTCAAGGCAGGGGTCTCCAGGATCCGTATATCCCAAGCAGAAACCATG | 831             |                   |
| Qy                    | 217    | GGACATCCCTTCGGGTCCTGGTGGCCACCCAGCATGTGGGTGCCATTATTGCGAAGGA     | 276             |                   |
| Db                    | 832    | TGATTTGGCTCTGGCCTGCTGGTTCCACCCTAATTTGTTGGAGCCATCATAGAAAGA      | 891             |                   |
| Qy                    | 277    | GGGGGCCACCATCGCGCAACATCAAAACAGACCCAGTCCCAAGATAGACGTGCATAGAA    | 336             |                   |
| Db                    | 892    | AGGTGCCACCATTCGGAAACATCAACCAAGACCCAGTCTAAATCGATGTCCACCGTAA     | 951             |                   |
| Qy                    | 337    | GGAGAAACGAGGTGCAGCTGAAAAGCCATCAGTGTGCATCCACCCCTGAGGGCTGCTC     | 396             |                   |
| Db                    | 952    | AGAAAAATGCGGGGGCTGCTGAGAAATCGATTACTCTCTCTACTCTCTGAAAGCACCTC    | 1011            |                   |
| Qy                    | 397    | CTCCGCTGTGAAGATGATCTTTGGAGATTATGCATAAAGAGGCTAAGACACCCAAACGGC   | 456             |                   |
| Db                    | 1012   | TGCGGCTTTGAAGTCTATTCTGGAGATTATGCAATAGGAAGCTCAAGATATAAAATTCAC   | 1071            |                   |
| Qy                    | 457    | TGACGAGGTTCCCTCGAAGATCCCTGGCCCATATAAATTTGTAGGGGCTCTCATTTGGCAA  | 516             |                   |
| Db                    | 1072   | AGAAGAGATCCCTTTGAAGATTTTAGCTCATAATAAATTTGTTGGAAGTCTATTGGTAA    | 1131            |                   |
| Qy                    | 517    | GGAGGACGGAACCTGAAAGAGGTAGAGCAAGATACCGAGACAAAAATACCAATCTCTC     | 576             |                   |
| Db                    | 1132   | AGAGGGAAGAATCTTAAAAAAATTTGACCAAGACACACACACTAAAAATCAGATATCTCC   | 1191            |                   |

|    |      |      |      |       |        |      |       |       |      |       |     |      |      |      |      |        |                     |                |                  |              |                 |             |              |     |      |               |                  |                |             |                |      |
|----|------|------|------|-------|--------|------|-------|-------|------|-------|-----|------|------|------|------|--------|---------------------|----------------|------------------|--------------|-----------------|-------------|--------------|-----|------|---------------|------------------|----------------|-------------|----------------|------|
| Qy | 577  | GTTC | AAGA | CTTAC | CCCTTT | TAC  | AACTG | TAG   | AGG  | ACAT  | CAC | TGT  | GA   | GGG  | CCAT | CGA    | 636                 |                |                  |              |                 |             |              |     |      |               |                  |                |             |                |      |
| Db | 1192 | ATTG | CAG  | GAAT  | TGAC   | GTGT | ATAAT | CCAG  | AACG | CAC   | TAT | TAC  | AGT  | TAA  | AGC  | CAATG  | TGA 1251            |                |                  |              |                 |             |              |     |      |               |                  |                |             |                |      |
| Qy | 637  | GAA  | TTG  | TG    | CAG    | GGCC | GAG   | CAG   | GAAT | ATAAT | TG  | AG   | AA   | GT   | TCGG | AGGCCT | TATGAGATGA 696      |                |                  |              |                 |             |              |     |      |               |                  |                |             |                |      |
| Db | 1252 | GAC  | ATG  | TG    | CCAA   | AGCT | TAG   | GAG   | GAG  | ATCAT | GA  | GA   | AAAT | CAGG | AGT  | CTTTAT | GAAATGA 1311        |                |                  |              |                 |             |              |     |      |               |                  |                |             |                |      |
| Qy | 697  | TG   | TG   | CT    | TG     | CCAT | GA    | ----- | GCT  | CT    | CAC | CT   | GA   | TCC  | TG   | CGCT   | CAACTGCTGTGATGG 750 |                |                  |              |                 |             |              |     |      |               |                  |                |             |                |      |
| Db | 1312 | TAT  | TG   | CT    | TCT    | TAAT | CTT   | CA    | AG   | CAAT  | TAA | TTCT | TG   | GA   | TAA  | CT     | GAA                 | CGCCTTGGG 1371 |                  |              |                 |             |              |     |      |               |                  |                |             |                |      |
| Qy | 751  | TC   | TTT  | CC    | AG     | CTT  | CA    | TG    | AG   | CG    | AG  | TCC  | CG   | CG   | CTT  | CC     | CAG                 | AGCGT          | TACTGGGGCTGC 810 |              |                 |             |              |     |      |               |                  |                |             |                |      |
| Db | 1372 | TC   | TG   | TT    | CC     | CA   | CC    | CA    | CTT  | CAG   | GG  | ATG  | CC   | AC   | CT   | CC     | AG                  | GG             | CCCCCTTCAG       | CCATGAC 1431 |                 |             |              |     |      |               |                  |                |             |                |      |
| Qy | 811  | TC   | CT   | A     | TAG    | CT   | CC    | TT    | AT   | TG    | CAG | GT   | CC   | CG   | AG   | AG     | AG                  | ATG            | GT               | CCAGGT       | GTTTATCCCCG 870 |             |              |     |      |               |                  |                |             |                |      |
| Db | 1432 | TC   | CT   | CC    | CT     | A    | CC    | CG    | AG   | TT    | TG  | AG   | CA   | T    | GA   | AA     | C                   | GG             | AG               | CTG          | TTCA            | TCCAGC 1491 |              |     |      |               |                  |                |             |                |      |
| Qy | 871  | CC   | AG   | CAG   | TGG    | GG   | CC    | AT    | CA   | TG    | CG  | AA   | AG   | GG   | CG   | CAG    | CA                  | CAT            | CA               | AA           | CAG             | CT          | CTCCCGTT 930 |     |      |               |                  |                |             |                |      |
| Db | 1492 | TC   | T    | AT    | CAG    | T    | CG    | T     | GC   | AT    | CA  | TG   | CG   | AA   | G    | AG     | GG                  | CC             | CAG              | CA           | AT              | CA          | AG           | CAG | CTT  | TCTCGCTT 1551 |                  |                |             |                |      |
| Qy | 931  | TG   | CAG  | CG    | CT     | CC   | AT    | CA    | AG   | ATT   | TG  | CA   | CC   | CG   | AAA  | CA     | CT                  | TG             | AT               | CT           | CC              | AA          | AG           | TT  | CG   | TAT           | TGTT 990         |                |             |                |      |
| Db | 1552 | TG   | CT   | GAG   | CTT    | CA   | ATT   | AA    | GA   | TT    | GT  | CT   | CA   | GG   | AA   | G      | CA                  | CA             | GA               | TG           | CT              | AA          | AG           | TG  | AG   | GA            | TGTT 1611        |                |             |                |      |
| Qy | 991  | TAT  | CA   | T     | CAC    | TG   | AC    | CG    | CA   | GG    | CC  | CA   | TT   | CA   | AG   | GT     | CAG                 | GG             | AG               | AT           | CT              | AT          | CG           | CA  | AA   | CT            | CGCAACT 1050     |                |             |                |      |
| Db | 1612 | GAT  | TAT  | CAC   | TG     | AC   | CA    | CG    | AG   | CT    | CAG | T    | CA   | AG   | GT   | CAG    | GG                  | AA             | GA               | AT           | TT              | TAT         | G            | GA  | AAAA | AT            | TATGGAAAAAT 1671 |                |             |                |      |
| Qy | 1051 | CA   | AG   | G     | GG     | AA   | CT    | TT    | T    | TG    | CT  | CC   | CA   | AG   | G    | AA     | GT                  | GA             | AG               | CT           | GG              | AG          | CC           | CA  | CA   | T             | AT               | CGTGT 1110     |             |                |      |
| Db | 1672 | TAA  | AG   | AA    | AA     | CT   | TT    | T     | TG   | AT    | CT  | T    | AA   | AG   | AG   | GT     | GA                  | AA             | CT               | T            | GA              | AG          | CT           | CA  | T    | AT            | CA               | GAGT 1731      |             |                |      |
| Qy | 1111 | GC   | AG   | CA    | T      | AG   | AG    | CT    | TG   | CG    | CG  | GT   | CA   | TT   | TG   | CA     | AA                  | AG             | GT               | TG           | GA              | AAAA        | CC           | GG  | T    | GA            | AC               | TTGCA 1170     |             |                |      |
| Db | 1732 | GCC  | AT   | CC    | TT     | TG   | CT    | GT    | CG   | CA    | AG  | AT   | TT   | TG   | GA   | AA     | G                   | AG             | G                | CA           | AA              | CC          | GG           | T   | GA   | AT            | CA               | AACTTCA 1791   |             |                |      |
| Qy | 1171 | GAA  | TTT  | G     | CG     | CAG  | CT    | CAG   | T    | GT    | GT  | AT   | CA   | AG   | AC   | CA     | CG                  | CC             | CT               | CA           | T               | AG          | AG           | CA  | CG   | CA            | CA               | CGCA 1230      |             |                |      |
| Db | 1792 | GAA  | TTT  | TG    | CA     | AG   | TG    | CA    | AG   | AT    | T   | GT   | T    | CC   | CT   | CG     | CA                  | CG             | AC               | CA           | CT              | GT          | AT           | GAG | AA   | T             | GACCA 1851       |                |             |                |      |
| Qy | 1231 | GG   | T    | AT    | CG     | T    | GA    | AA    | AT   | CA    | T   | CG   | GA   | AT   | TT   | CT     | TAT                 | G              | CG               | CA           | T               | AG          | T            | GG  | CT   | CA            | CG               | GAAGATCCG 1290 |             |                |      |
| Db | 1852 | AG   | T    | GG    | T      | TG   | CA    | AA    | TA   | CT    | GT  | CA   | TT   | CA   | T    | AT     | GT                  | T              | G                | CC           | AG              | TT          | T            | G   | CC   | AG            | T                | T              | GCCAGGTTGCC | CCAGAGAAAAATTC | 1911 |
| Qy | 1291 | AG   | CA   | T     | CT     | CG   | CC    | AG    | TT   | AA    | G   | CAG  | CAG  | CA   | T    | CA     | GA                  | G              | 1327</           |              |                 |             |              |     |      |               |                  |                |             |                |      |

```

RESULT 12
US-09-480-884A-175
; Sequence 175, Application US/09480884A
; Patent No. 6482597
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121-455C6
; CURRENT APPLICATION NUMBER: US/09/480,884A
; CURRENT FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 330
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 175
; LENGTH: 4181

```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(4181)
; OTHER INFORMATION: n=A,T,C or G
US-09-480-884A-175

Query Match      34.0%; Score 580.6; DB 4; Length 4181;
Best Local Similarity 66.1%; Pred. No. 1.2e-155;
Matches 857; Conservative 0; Mismatches 434; Indels 6; Gaps 1;

QY 37 AGCCATCATGAAGCTGAATGGCCACCACTGTGGAGAACCAATCCCTCGAAGGTCTCTACAT 96
Db      |||||
QY 652 AGCACTAGACAACTGAATGGATTTCAGTTAGAGAAATTCACCTTGAAGTAGCCTATAT 711
Db      |||||
QY 97 CCCCATGAGCAGATAGCACAGGACCTGAGATGGGCGCCGAGGGGCTTGGCTCTCG 156
Db      |||||
QY 712 CCTGATGAATGGCGCCGACAAACCCCTTGCAGCAGCCCGAGGTCCCGGGGGCT 771
Db      |||||
QY 157 GGGTCAGCCCGCCAGGGCTCACCTGTGGCAGCGGGGCCCCAGCCAGCAGCAGCAAGT 216
Db      |||||
QY 772 TGGGAGAGGGGCTCCTCAAGCGAGGGTCTCCAGGATCCGTATCCAGCAGAAACCATG 831
Db      |||||
QY 217 GGACATCCCTTGGCTCTGGTCCGCCACCGAGTATGTGGGTGCCATTAATTTGCAAGGA 276
Db      |||||
QY 832 TGAATTTGGCTCTGGCGCTCTGGTCTCCACCAATTTGTGGAGGCCATCATAGGAAAGA 891
Db      |||||
QY 277 GGGGCCCCACCATCCGCAACATCACAAAACAGACCCAGTCCCAAGATAGACGTGCATAGAA 336
Db      |||||
QY 892 AGTGCCACCATTCGGAACATCACCAACACAGCCAGTCTAAATCGATGTCCACCGTAA 951
Db      |||||
QY 337 GGAGAACCGAGTGCAGCTGAAAAAGCCATCAGTGTGCATCCACCCCTGAGGCTCTC 396
Db      |||||
QY 952 AGAAAAATCGGGGCTGTGAGAGTGTACTATCTCTACTCTGCTGAGGCACTC 1011
Db      |||||
QY 397 CTCCTCTGTAGATGATCTTGGAGTATGCATTAAGAGGCTAAGGACACCAAAACGGC 456
Db      |||||
QY 1012 TCGGCTTGTAGTCTTCTGGAGATATGATCAAGGAAGTCAAGATATAAATTCAC 1071
Db      |||||
QY 457 TGACGAGGTCTCCCTGAAGATCTGGCCCAATAAATTTGTAGGGCTCTCATTTGCCAA 516
Db      |||||
QY 1072 AGAAGAGATCCCTTGAAGATTTAGCTCATAATAACTTTGTGGAAGTCTTATTGGTAA 1131
Db      |||||
QY 517 GGAAGCAGGAACCTTGAAGAGGTAGACCAAGATACCGAGCAAAATCAACCTCTCTC 576
Db      |||||
QY 1132 AGAAGGAAGAATCTTAAAAAATTTAGCAAGACACACACATAAAATCAGATATCTCC 1191
Db      |||||
QY 577 GTTCAAGACCTTACCCTTTACACCCCTGAGAGGACCATCACTGTGAAGGGGGCCATCGA 636
Db      |||||
QY 1192 ATTGAGGAATGACCTGTATATCCAGACGCACTATTACGTTAAGGCAATGTGA 1251
Db      |||||
QY 637 GAATTTGTCAGGGCCGAGAGGAATAATGAAGAACTTCGGAGGCTTATGAGATGA 696
Db      |||||
QY 1252 GACATGTGCAAAAGCTGAGGAGGATCATGAAGAAATCAGGAGTCTTATGAAATGA 1311
Db      |||||
QY 697 TGTGGCTGCCATGA-----GCTCTCACTGATCCCTGGCTGAACTGGCTGTGAGG 750
Db      |||||
QY 1312 TATTGCTTCTATGAATCTTCAAGCACTTTAATTCCTGATTAATCTGAACGCTTGGG 1371
Db      |||||
QY 751 TCTTTTCCAGCTTCATCAGCGCAGTCCCGCGCTCCAGCAGCGTTCAGTGGGCTGC 810
Db      |||||
QY 1372 TCTGTTCCACCCACTTCAGGAGTCCACCTCCACCTCAGGCGCCCTTCAGCCATGAC 1431
Db      |||||
QY 811 TCCCTATAGCTCTTTATGAGGCTCCCGAGCAGAGATGTGCGAGTGTATTATCCCGC 870
Db      |||||
QY 1432 TCTCTCCCTACCGGAGTTGAGCAATCAGAAACGGAGACTGTTCATCAGTTTATCCCGC 1491
Db      |||||
QY 871 CCAGGAGTGGCGGCATCATCGGCAAGAGGGGCGAGCAGCATCAACAGCTCTCCCGGT 930
Db      |||||
QY 1492 TCTATCAGTCCGTGCGCATCATCGGCAAGCGGCGCCAGCAGCATCAAGCAGCTTCTCGCT 1551
Db      |||||
QY 931 TGCCAGCGCTCCATCAAGATTGCAACACCCAGAACCTGACTCCAAAGTTCGTATGGT 990
Db      |||||

1552 TGCTGAGCTTCAATTAAGATTGCTCCAGCGGAAGCACAGATGCTAAAGTGAGGATGGT 1611
991 TATCATCTGACCGCCAGAGGCCCAATTCAGGCTCAGGGAAGAATCTATGGCAAACT 1050
1612 GATTATCACTGGACCACAGAGGCTCAGTTCAAGGCTCAGGGAAGAATTTATGGAAAAAT 1671
1051 CAAGGAGGAGAACTTCTTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACCCACATACGTGT 1110
1672 TAAAGAGAAACTTTGTTAGTCTTAAAGAGAGGTGAACCTTGAAGCTCATATCAGAGT 1731
1111 GCGAGATCAGCAGCTGCGCGGTCAATTGGCNAAGGTGAAAAACGGTGAACGAGTTGCA 1170
1732 GGCATCTTTGCTGCTGGCAGAGTTATTGGAAAAAGGCGCAAAACGGTGAATGAACCTTCA 1791
1171 GAATTTGACGGCAGCTGAGGTGTAGTACCAAGAGACACAGACCCCTGATGAGAACGACA 1230
1792 GAATTTGTCAGTGAGAGTTGTTGCTCCTGTCACAGACACCTGATGAGNATGACCA 1851
1231 GGTCACTCGTGAATAATCATCGGACATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCGG 1290
1852 AGTGGTTGTCAAAATAACTGCTCACTTCTATGCTTGCAGGTTGCCAGAGAAAAATTCA 1911
1291 AGACATCTCGGCCAGGTTAAGCAGCAGCATCAGAG 1327
1912 GGAATTTCTGACTCAGGTTAAGCAGCAGCACCAACAG 1948

RESULT 13
US-09-542-615A-175
; Sequence 175, Application US/09542615A
; Patent No. 6518256
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; FILE REFERENCE: 210121.455C8
; CURRENT APPLICATION NUMBER: US/09/542,615A
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 175
; LENGTH: 4181
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (3347)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3502)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3506)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3520)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3538)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3549)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3646)
; OTHER INFORMATION: n=A,T,C or G
```

```
; LOCATION: (3940)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3968)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3974)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (4036)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (4056)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (4062)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (4080)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (4088)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (4115)
; OTHER INFORMATION: n=A,T,C or G
; OTHER INFORMATION: n=A,T,C or G
US-09-542-615A-175

Query Match      34.0%; Score 580.6; DB 4; Length 4181;
Best Local Similarity 56.1%; Pred. No. 1.2e-155;
Matches 857; Conservative 0; Mismatches 434; Indels 6; Gaps 1;

QY 37 AGCATATGAGCTGAATGCCACAGTTCGAGAACCATGCCCTGAGGCTCCTACAT 96
Db 652 AGCACTAGACAACTGAATTCAGTTAGAGATTTCACTTGAAGTAGCCTATAT 711
QY 97 CCCCAGATGAGCAGATAGACACAGGACCTGAGAAATGGCGCGAGGGGCTTTGGCTCTCG 156
Db 712 CCCTGATGAATGGCGCGCCAGCAAAACCCCTTGACGAGCCCCGAGGTGCGCGGGCT 771
QY 157 GGGTCAGCCCGCCAGGAGCTCACCTGTGGCAGCGGGGCCCCAGCAGCAGCAAGT 216
Db 772 TGGGCAGAGGGGCTCCTCAAGCAGGGGTCTCCAGGATCCGTATCCAGCAGAAACCATG 831
QY 217 GGACATCCCTTCCGCTCCTGGTCCACCCAGTATGTGGTGCCATTATTGGCAAGGA 276
Db 832 TGATTTGGCTCTGCGCTCTGCTGTTCCACCCCAATTTGTTGGACCATCATAGGAAGA 891
QY 277 GGGGCGCCACATCCGCAACATCAAAAAGACAGCCAGTCCAGATAGACGTGCATAGAA 336
Db 892 AGGTGCCACCATTCGGAACATCAACAAACAGACCCAGTCTAAAATCGATGTCCACCGTAA 951
QY 337 GGAGAACGAGGTGCAGCTGAAAAGCCATCAGTGTGCATCCACCCCTGAGGGCTGCTC 396
Db 952 AGAAATGGGGGGTGTGAGAGTGAATTAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1011
QY 397 CTCCGCTTTGAAGATGATCTTTGAGATATGATCAATAAGAGGCTTAAGGACACCAAAACGGC 456
Db 1012 TCGCGCTTTGAAGTCTATTTCTGGAGATTATGATAAGGAAGCTCAAGATATAAAATTCAC 1071
QY 457 TGACAGAGTTCCTTCAGATCTCTGGCCATAATTAATCTTTAGGGGCTCTCATTTGGCAA 516
Db 1072 AGAAGAGATCCCTTTGAAGATTTTAGCTCATAATAAATTTTGTGGAGCTCTTATTTGGTAA 1131
QY 517 GGAAGGACGGAACCTTGAAGAAGGTAGAGACAGATACCGAGACAAAATCACCATCTCTCTC 576
Db 1132 AGAAGGAAGAATCTTAAATAATTTGAGCAAGACACACACTAAATCAGATATCTCC 1191
QY 577 GTTGCAGACCTTACCTTTTACACCTGTAGAGGACCATCATCTGTGAAGGGGGCCATCGA 636
Db 1192 ATTGCAGGAATTCAGCTGTATAATCCAGAACGCACTATTACAGTTAAAGGCAATGTGA 1251
QY 637 GAATTTGTTGCAGGGCGGACGAGCAATAATAGAAAGTTTCGGGAGGGCTATGAGATGA 696
```

```
Db 1252 GACATGTGCCAAAGCTGAGGAGGAGATCATGAAGAAAATCAGGAGTCTTTATGAAATGA 1311
QY 697 TGTGGCTGCCATGA-----GCTCTCACCTGTATCCCTGGCCCTGAACCTGGCTGCTTAGG 750
Db 1312 TATTGCTTCTATGAATCTTCAAGCACATTTAATCTCTGATTAATCTGAACGCTTGG 1371
QY 751 TCTTTTCCAGCTTCTATCCAGCGCAGTCCCGCGCGCTCCACAGAGCGTTACTTGGGCTGC 810
Db 1372 TCTGTTCCCAACCACTTCAGGGATGCCAGCTCCCACTCAGGGCCCCCTTCAGGCCATGAC 1431
QY 811 TCCCTATAGCTCCTTTATGCAGGCTCCCGAGCAGGAGATGGTGCAGGTGTTTATCCCCGC 870
Db 1432 TCTCCCTAGCCGACGTTTGAACAATCAGAACGAGACTGTTCATCAGTTTATCCAGC 1491
QY 871 CCAGGAGTGGCGCCATCATCGGCAAGAGGGGAGCAGACATCAAAACAGCTCTCCCGGT 930
Db 1492 TCTATCAGTCGCTGCCATCATCGGCAAGCAGGCGCCAGCACATCAAGCAGCTTTCTCGCT 1551
QY 931 TGCAGGCGCTCCATCAAGATTGCACACCCGAAACCTGACTCCAAAGTTCGTATGTT 990
Db 1552 TGCTGAGCTTCAATTAAGATTGCTCCAGCGGAAGCACCAGATGCTAAAGTGAGGATGGT 1611
QY 991 TATCATCACTGGACCGCCAGAGGCCCAATTTCAAGGCTCAGGGAAGAAATCTATGGCAAACT 1050
Db 1612 GATTATCACTGGACCAACAGAGGCTCAGTTCAGGCTCAGGGAAGAAATTTATGGAATAAT 1671
QY 1051 CAAGGAGGAACTTCTTTGGTCCCAAGGAGGAGTGAAGCTGGAGACCCACATACGTGT 1110
Db 1672 TAAAGAAGAAAATCTTTGTAGTCTTAAAGAAGAGGTGAAACTTTGAAGCTCATATCAGAGT 1731
QY 1111 GCCAGCATCAGCAGCTGGCGGCTCATTTGCCAAAGGTGAAATAACCGTGAACGAGTTGCA 1170
Db 1732 GCATCTCTTGTCTGTCGAGAGTTATTGGAAAGAGGCGCAAAACCGTGAATGAATTCAT 1791
QY 1171 GAATTTAGCGGACGTGAGTGGTAGTACCAAGAGACCAAGACCCCTGATGAGAACCAACA 1230
Db 1792 GAATTTGTCAAGTGCAGAAAGTTGTTGTCTCTGTGACCAAGACACCTGATGAGAAATGACCA 1851
QY 1231 GGTCACTGTGAATCATCGACATTTCTATGCCACTCAGATGGCTCAACGGAAGATCCG 1290
Db 1852 AGTGGTTGTCAAATAACTGGTCACTTCTATGTTGCCAGTTGCCAGAGAAAATTCAT 1911
QY 1291 AGACATCTCGGCCCGAGGTTAAGCAGCAGCATCAGAAAG 1327
Db 1912 GGAAATCTGACTCAGGTAAGCAGCAGCAACCAACACAG 1948

RESULT 14
US-09-606-421B-175
; Sequence 175, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 175
; LENGTH: 4181
; TYPE: DNA
; ORGANISM: Homo sapiens
```







QY 1171 GAATTGACGCGAGCTGAGGTGGTAGTACCAAGAGAGACGAGACCCCTGATGAGAGACCA 1230  
Db |||||  
QY 1792 GAATTGTCAGTCAGAGAGTGTGTCCTCGTGACAGACACCTGATGAGATGACCA 1851  
Db |||||  
QY 1231 GGTATCGTGAATAATCATCGACATTTCTATGCCAGTCAGATGCTCAACGGAAGATCCG 1290  
Db |||||  
QY 1852 AGTGGTTGCAAAATAACTGTCACCTCTATGCTTGCAGGTTGCCAGAGAAAAATTCA 1911  
Db |||||  
QY 1291 AGACATCTGCCGCCAGGTTRAGCAGCAGCATCAGAG 1327  
Db |||||  
QY 1912 GGAATTCGACTCAGGTAAAGCAGCACCAACACAG 1948  
Db |||||  
RESULT 15  
US-09-061-709-8  
; Sequence 8, Application US/09061709B  
; Patent No.. 6297364  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Yao-Tsung  
; APPLICANT: Gure, Ali  
; APPLICANT: Tsang, Solam  
; APPLICANT: Stockert, Elisabeth  
; APPLICANT: Jager, Elke  
; APPLICANT: Knuth, Alexander  
; APPLICANT: Old, Lloyd J.  
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated  
; FILE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof  
; FILE REFERENCE: LUD 5538  
; CURRENT APPLICATION NUMBER: US/09/061,709B  
; CURRENT FILING DATE: 1998-04-17  
; NUMBER OF SEQ ID NOS: 8  
; SEQ ID NO 8  
; LENGTH: 3283  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
US-09-061-709-8  
Query Match 29.5%; Score 504; DB 3; Length 3283;  
Best Local Similarity 63.6%; Pred. No. 8.4e-134;  
Matches 849; Conservative 0; Mismatches 425; Indels 60; Gaps 3;  
QY 37 AGCCATCATGATGATGATGCCACCATGTTGGAGNACCATGCCCTGAAGTCTCCTACAT 96  
Db 474 AGCCATGAGAGAGCTTAAGCGGGCATCAGTTTGGAGAACTACTCTCTCAAGATTTCCTACAT 533  
QY 97 CCCCAGTACGAGATAGACACAGGACCTGAGATGGGCGCGCGGGGGCTTTGGCTCTCG 156  
Db 534 CCGGATGA-----AGAGGTGAGCTCCCTTCGCCCTCAGCGAGCCAGCGTGG 584  
QY 157 GGGTCAGCCCGCAGAGGCTCCTGTTGGAGCGGGGGGCCCGCCAGCAGCAGCAAGT 216  
Db 585 GGACCACTTTCCCGGGAGCAAGGCCACGCCCTCGGGGGCACTTTCTCAGGCGCAGACAGAT 644  
QY 217 GGACATCCCTTCCGCTCTGGTGGTCCACCCAGTATGTTGGTGCCATTTATGGCAAGGA 276  
Db 645 TGAATTCCTGCGATGCTGTTGCCACCCAGCTTTGTTGGTGCCATCATCGAAGGA 704  
QY 277 GGGGGCCACCATCCGCAACATCAAAACAGACCCAGTCCAAAGATAGACGTCATAGGAA 336  
Db 705 GGGGTTGACCAATAAAGAACATCACTAAGCAGACCCAGTCCCGGGTAGATATCCATAGAAA 764  
QY 337 GGAGAACGAGTGCAGCTGAAGAGCCATCAGTGTGCATCCACCCCTGAGGGCTCTC 396  
Db 765 AGAGAACTCTGGAGTGCAGAGAGCCGTGTCACCATCCATGCCACCCAGAGGGGACTTC 824  
QY 397 CTCGCTTGTAGATGATCTTTGGAGATTATGCATAAAGAGGCTAAGGACACCAAAACGGC 456  
Db 825 TGAAGCATGCCGATGATCTTTGAAATCATCGAAGAGAGCCAGATGAGACCAACTAGC 884  
QY 457 TGACAGAGTTCCTCGAAGATCCTGGCCCAATAAATCTTTGTAGGGCGTCTCATTTGGCAA 516  
Db 885 CGAAGAGATTCCTCTGAAAATCTTTGGCACAAATGGTGTGGTGGAAAGACTGATTTGAAA 944

Search completed: July 13, 2004, 21:42:51  
Job time : 141 secs

QY 517 GGHAGACGGAACCTGAAGAGGTAGAGCAAGATACCAGAGACAAATAATCACCATCTCCTC 576  
Db |||||  
QY 945 AGAAGCGAGAAATTTGAAGAAATTTGAACATGAAACAGGACCAAGATAAACAATCTCATC 1004  
Db |||||  
QY 577 GTTGAAGACCTTTACCCCTTTTCAACCCCTGAGAGGACCATCACTGTGAAGGGGCCATCGA 636  
Db |||||  
QY 1005 TTTGCAAGATTTCAGCATATCAACCCGGAAGAACCATCACTGTGAAGGGCACAGTTGA 1064  
Db |||||  
QY 637 GAATTTGTCAGGGCCGAGCAGCAAAATATGAAGAAAGTTGCGGAGGCTTATGAGAAATGA 696  
Db |||||  
QY 1065 GGCCTGTGCCAGTGTGAGATAGAGATTATGAAGAAGTGTGCGTAGGCGCTTTGAAATGA 1124  
Db |||||  
QY 697 TGTGGCTGCGATGAGCTCTCACTGATCCCTGGCCTGAACTGGCTGTAGGCTTTT 756  
Db |||||  
QY 1125 TATGCTGGTGTAAACCCCA-----CTCCGGATACTTCTCC 1161  
QY 757 CCGAGTTTCACTCAGCGCAGTCCCGCGCTCCAGCAGCTTCTGAGGCTGCTCCCTA 816  
Db |||||  
QY 1162 AGCCTGTAGCCCATCACCAGTTTGGCGGTTTCCCGCATCACT----- 1207  
QY 817 TAGCTCTTTATGCGAGCTCCCGAGCAGGAGATGTGCGAGGTGTTTATCCCGCCAGGC 876  
Db |||||  
QY 1208 -----CTTATCCAGAGCAGGAGATTGTGAATCTCTTCATCCCAACCCAGGC 1253  
QY 877 AGTGGCGGCATCATCGCAAGAAAGGGGAGCAGCAGCATCAAAACAGCTCTCCCGGTTGGCCAG 936  
Db |||||  
QY 1254 TGTGGCGGCATCATCGGAAGAGGGGGCACACATCAAAACAGCTGGCGAGATTGCGCGG 1313  
QY 937 CGCTCTCATCAAGATTGACACCCGAGAAACACTGACTCCAAAGTTCGATGTTGTTATCAT 996  
Db |||||  
QY 1314 AGCCTCTATCAAGATTGCCCCCTGCGAAGGCCGAGACGTCAGCGAAAGGATGTCATCAT 1373  
QY 997 CACTGGACCGCAGAGGCGCCAAATTCAGAGCTCAGGGAAGAAATCTATGGCAAACTCAAGGA 1056  
Db |||||  
QY 1374 CACCGGGCCACCGGAGCCAGTTCAAGGCCAGGACGATCTTTGGGAAACTGAAAGA 1433  
QY 1057 GGAGAACTTTTGGTCCCAAGGAGAGTGAAGCTGGAGACCCACATACGTGTGCCAGC 1116  
Db |||||  
QY 1434 GGAAAACTTTCTTTAAACCCCAAGAAAGTGAAGCTGGAAAGGCCATATCAGAGTGGCCCTC 1493  
QY 1117 ATCAGCAGCTGGCGGGTCTATTTGGCAAAAGGTGGAAGAACCGTGAACGAGTTTCAGAAATTT 1176  
Db |||||  
QY 1494 TTCACAGCTGGCGGGTGAITTTGGCAAAAGGTGGCAAGACCGTGAACGAACTGCAGAACTT 1553  
QY 1177 GACGGCAGCTGAGGTGGTAGTACCAAGAGACAGACCCCTGTATGAGAACGACAGGTCAT 1236  
Db |||||  
QY 1554 AACCACTGCAGAAATCATCGTGCCTCGTGACCAACGCCAGATGAAATGAGGAAGTAT 1613  
QY 1237 CGTGAATAATCATCGGACATTTCTATGCCAGTCAAGTGGCTCAACGGAAGATCGAGACAT 1296  
Db |||||  
QY 1614 CGTCAAGATTATCGGGCACTTTCTTTGCTAGCCAGACTGCACAGCGCAAGATCAGGGAAT 1673  
QY 1297 CTTGGGCCAGGTTAAGCAGCAGCATCAGAAAGGAGACAGATTAACAGGCGCCAGGACGAG 1356  
Db |||||  
QY 1674 TGTACAAAGGTGAAAGCAGCAGGAGCAGGAGGAGAAATACCTCAGGAGTGCCTCTCAGCGCAG 1733  
QY 1357 GAAGTACAGGCC 1370  
Db |||||  
QY 1734 CAAGTCAGGCTGCC 1747



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 13, 2004, 14:54:07 ; Search time 6722 Seconds  
(without alignments)  
11013.077 Million cell updates/sec

Title: US-09-270-437D-5  
Perfect score: 1708  
Sequence: 1 agggacgtgcgcacgcc.....atttccttcaggttttaaaa 1708

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_hg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_ov.\*
- 22: em\_pat.\*
- 23: em\_ph.\*
- 24: em\_pl.\*
- 25: em\_ro.\*
- 26: em\_sts.\*
- 27: em\_un.\*
- 28: em\_vi.\*
- 29: em\_vl.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_mam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match % | Length | DB | ID       | Description        |
|------------|--------|---------------|--------|----|----------|--------------------|
| 1          | 1703   | 99.7          | 1708   | 6  | AR171864 | AR171864 Sequence  |
| 2          | 1703   | 99.7          | 1708   | 6  | AR343073 | AR343073 Sequence  |
| 3          | 1703   | 99.7          | 1708   | 6  | BD209924 | BD209924 Isolated  |
| 4          | 1669.2 | 97.7          | 1946   | 6  | AR171866 | AR171866 Sequence  |
| 5          | 1669.2 | 97.7          | 1946   | 6  | AR343075 | AR343075 Sequence  |
| 6          | 1669.2 | 97.7          | 1946   | 6  | BD209926 | BD209926 Isolated  |
| 7          | 1653.2 | 96.8          | 2130   | 9  | AF117106 | AF117106 Homo sapi |
| 8          | 1631.8 | 95.5          | 2381   | 9  | AF198254 | AF198254 Homo sapi |
| 9          | 1389   | 81.3          | 2780   | 6  | BD127811 | BD127811 Primer fo |
| 10         | 1389   | 81.3          | 2780   | 9  | AK074915 | AK074915 Homo sapi |
| 11         | 1165   | 68.2          | 2224   | 6  | AR160244 | AR160244 Sequence  |
| 12         | 1162.6 | 68.1          | 2223   | 10 | AF061569 | AF061569 Mus muscu |
| 13         | 1161.6 | 68.1          | 2244   | 10 | BC051879 | BC051879 Mus muscu |
| 14         | 1112.4 | 65.1          | 2105   | 10 | AF541940 | AF541940 Rattus no |
| 15         | 914.8  | 53.6          | 2021   | 5  | AF026527 | AF026527 Gallus ga |
| 16         | 732.8  | 42.9          | 833    | 6  | BD124998 | BD124998 Primer fo |
| 17         | 732.8  | 42.9          | 833    | 6  | BD126684 | BD126684 Primer fo |
| 18         | 627    | 36.7          | 1740   | 10 | AB046173 | AB046173 Mus muscu |
| 19         | 627    | 36.7          | 1740   | 10 | BC045138 | BC045138 Mus muscu |
| 20         | 627    | 36.7          | 4030   | 10 | BC049082 | BC049082 Mus muscu |
| 21         | 588    | 34.4          | 2011   | 5  | AF161270 | AF161270 Danio rer |
| 22         | 588    | 34.4          | 2612   | 5  | BC045873 | BC045873 Danio rer |
| 23         | 582.2  | 34.1          | 1740   | 6  | BD275902 | BD275902 COMPOUNDS |
| 24         | 582.2  | 34.1          | 1740   | 6  | AR220687 | AR220687 Sequence  |
| 25         | 582.2  | 34.1          | 1740   | 6  | AR281251 | AR281251 Sequence  |
| 26         | 582.2  | 34.1          | 1740   | 6  | AX365954 | AX365954 Sequence  |
| 27         | 582.2  | 34.1          | 1743   | 6  | AX366054 | AX366054 Sequence  |
| 28         | 582.2  | 34.1          | 1743   | 6  | AX366057 | AX366057 Sequence  |
| 29         | 582.2  | 34.1          | 1764   | 6  | AX366035 | AX366035 Sequence  |
| 30         | 582.2  | 34.1          | 4171   | 9  | AF117108 | AF117108 Homo sapi |
| 31         | 580.6  | 34.0          | 4155   | 9  | HSU76705 | U76705 Human putat |
| 32         | 580.6  | 34.0          | 4159   | 6  | AR171863 | AR171863 Sequence  |
| 33         | 580.6  | 34.0          | 4159   | 6  | AR343072 | AR343072 Sequence  |
| 34         | 580.6  | 34.0          | 4159   | 6  | BD209923 | BD209923 Isolated  |
| 35         | 580.6  | 34.0          | 4181   | 6  | BD275767 | BD275767 COMPOUNDS |
| 36         | 580.6  | 34.0          | 4181   | 6  | AR220552 | AR220552 Sequence  |
| 37         | 580.6  | 34.0          | 4181   | 6  | AR255546 | AR255546 Sequence  |
| 38         | 580.6  | 34.0          | 4181   | 6  | AR281116 | AR281116 Sequence  |
| 39         | 580.6  | 34.0          | 4181   | 6  | AX333233 | AX333233 Sequence  |
| 40         | 580.6  | 34.0          | 4181   | 6  | AX365782 | AX365782 Sequence  |
| 41         | 580.6  | 34.0          | 4181   | 9  | HSU97188 | U97188 Homo sapien |
| 42         | 580.6  | 34.0          | 4602   | 6  | AX397963 | AX397963 Sequence  |
| 43         | 571.8  | 33.5          | 2161   | 5  | AF055923 | AF055923 Xenopus l |
| 44         | 571.8  | 33.5          | 2405   | 5  | AF064634 | AF064634 Xenopus l |
| 45         | 562.6  | 32.9          | 2031   | 5  | BC057700 | BC057700 Xenopus l |

ALIGNMENTS

RESULT 1  
AR171864  
LOCUS AR171864 1708 bp DNA linear PAT 17-DEC-2001  
DEFINITION Sequence 5 from patent US 6297364.  
ACCESSION AR171864  
VERSION AR171864.1 GI:17910814  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
1 (bases 1 to 1708)  
AUTHORS Chen,Y.-T., Gure,A., Tsang,S., Stockert,E., Jager,E., Alexander,K.  
and Old,L.J.  
TITLE Isolated nucleic acid molecule encoding cancer associated antigen,  
the antigen itself, and uses thereof

|                            |   |  |                            |
|----------------------------|---|--|----------------------------|
| JOURNAL                    | Patent: US 6297364-A 5 02-OCT-2001;   | Location/Qualifiers  |                            |
| FEATURES                   | 1. .1708  | /organism="unknown"  |                            |
| source                     | /mol_type="unassigned DNA"  |  |                            |
| ORIGIN                     |   |  |                            |
| Query Match                | 99.7%; Score 1703; DB 6; Length 1708;   |  |                            |
| Best Local Similarity      | 100.0%; Pred. No. 0;  |  |                            |
| Matches 1708; Conservative | 0; Mismatches   | 0; Indels  | 0; Gaps                    |
| 0;                         |   |  |                            |
| Qy                         | 1   | AGGAGCGTGGCGGACCGCCCGCCAGTTTACCCCGGGAGCCATCATGAAGCTGAATGCCCA | 60                         |
| Db                         | 1   | AGGAGCGTGGCGGACCGCCCGCCAGTTTACCCCGGGAGCCATCATGAAGCTGAATGCCCA | 60                         |
| Qy                         | 61  | CCAGTTGAGAACCATGCCCTGAAGTCTCTACATCCCGATGAGCAGATAGCAGGG       | 120                        |
| Db                         | 61  | CCAGTTGAGAACCATGCCCTGAAGTCTCTACATCCCGATGAGCAGATAGCAGGG       | 120                        |
| Qy                         | 121   | ACCTGAGAAATGGGCGCGGAGGGCTTTGGCTCTCGGGGTGAGCCCGCCAGGGCTCAC    | 180                        |
| Db                         | 121   | ACCTGAGAAATGGGCGCGGAGGGCTTTGGCTCTCGGGGTGAGCCCGCCAGGGCTCAC    | 180                        |
| Qy                         | 181   | TGTGGCAGCGGGGCGCCAGCAAGCAGCAGCAAGTGGACATTCCTTGGCTCTGGT       | 240                        |
| Db                         | 181   | TGTGGCAGCGGGGCGCCAGCAAGCAGCAGCAAGTGGACATTCCTTGGCTCTGGT       | 240                        |
| Qy                         | 241   | GCCACACAGTATGTGGTGGCTTATTTGGCAAGGAGGGGCCACCATCCGACATCAC      | 300                        |
| Db                         | 241   | GCCACACAGTATGTGGTGGCTTATTTGGCAAGGAGGGGCCACCATCCGACATCAC      | 300                        |
| Qy                         | 301   | AAAACAGACCCAGTCCAGATAGACGTGATAGGAGGAGGAGGAGGAGGAGGAGGAGG     | 360                        |
| Db                         | 301   | AAAACAGACCCAGTCCAGATAGACGTGATAGGAGGAGGAGGAGGAGGAGGAGGAGG     | 360                        |
| Qy                         | 361   | AGCCATCAGTGTGACTCCACCCCTGAGGGTGTCTCTCCGCTTGAAGTATCTTGA       | 420                        |
| Db                         | 361   | AGCCATCAGTGTGACTCCACCCCTGAGGGTGTCTCTCCGCTTGAAGTATCTTGA       | 420                        |
| Qy                         | 421   | GATTATGATATAAGAGGTAAAGGACACCAAAACGGCTGACGAGGTTCCTTGAAGATCT   | 480                        |
| Db                         | 421   | GATTATGATATAAGAGGTAAAGGACACCAAAACGGCTGACGAGGTTCCTTGAAGATCT   | 480                        |
| Qy                         | 481   | GGCCATTAATACCTTTGTAGGGGTCTCATTTGGCAAGGAGGAGGAGGAGGAGGAGG     | 540                        |
| Db                         | 481   | GGCCATTAATACCTTTGTAGGGGTCTCATTTGGCAAGGAGGAGGAGGAGGAGGAGG     | 540                        |
| Qy                         | 541   | AGAGCAAGATACCGAGACAAAATACCATCTCTCTGTTGCAAGACCTTACCCCTTACAA   | 600                        |
| Db                         | 541   | AGAGCAAGATACCGAGACAAAATACCATCTCTCTGTTGCAAGACCTTACCCCTTACAA   | 600                        |
| Qy                         | 601   | CCCTGAGAGGACCATCATCTGTGAAGGGGCGCATCGAGAAATTTGTTGCGGGCCGAGCAG | 660                        |
| Db                         | 601   | CCCTGAGAGGACCATCATCTGTGAAGGGGCGCATCGAGAAATTTGTTGCGGGCCGAGCAG | 660                        |
| Qy                         | 661   | AAATATGAAGAGTTCGGAGGGCTATGAGAAATGATGGCTGCTGCTGCTGCTGCTGCT    | 720                        |
| Db                         | 661   | AAATATGAAGAGTTCGGAGGGCTATGAGAAATGATGGCTGCTGCTGCTGCTGCTGCT    | 720                        |
| Qy                         | 721   | GATCCCTGGCCTGAACCTGGTGTAGGTCTTTTCCAGCTTATCCAGCGCAGTCCC       | 780                        |
| Db                         | 721   | GATCCCTGGCCTGAACCTGGTGTAGGTCTTTTCCAGCTTATCCAGCGCAGTCCC       | 780                        |
| Qy                         | 781   | GGCGCTCCAGCAGCTTACCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT   | 840                        |
| Db                         | 781   | GGCGCTCCAGCAGCTTACCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT   | 840                        |
| Qy                         | 841   | CGAGGAGATGTGCAAGTGTATATCCCGCCAGCAGTGGCGCCCATCATCGGCAAGAA     | 900                        |
| Db                         | 841   | CGAGGAGATGTGCAAGTGTATATCCCGCCAGCAGTGGCGCCCATCATCGGCAAGAA     | 900                        |
| Qy                         | 901   | GGGGCAGCAGCATCAAGAGCTCTCCGGTTGGCGAGCGCTTCAAGTTGACACACC       | 960                        |
| Db                         | 901   | GGGGCAGCAGCATCAAGAGCTCTCCGGTTGGCGAGCGCTTCAAGTTGACACACC       | 960                        |
| Qy                         | 961   | CGAAACACCTGACTCCAAAGTTCGATGTTTATCATCTGAGCGCGAGAGGCCCAATT     | 1020                       |
| Db                         | 961   | CGAAACACCTGACTCCAAAGTTCGATGTTTATCATCTGAGCGCGAGAGGCCCAATT     | 1020                       |
| Qy                         | 1021  | CAAGGCTCAGGAGAAATCTATGGCAAACTCAAGGAGGAGAACTTTTGTGTCACAGGA    | 1080                       |
| Db                         | 1021  | CAAGGCTCAGGAGAAATCTATGGCAAACTCAAGGAGGAGAACTTTTGTGTCACAGGA    | 1080                       |
| Qy                         | 1081  | GGAAGTGAAGTGGAGACCCCATACAGTGTGCCAGCATCAGCAGTGGCCGGGTCAITGG   | 1140                       |
| Db                         | 1081  | GGAAGTGAAGTGGAGACCCCATACAGTGTGCCAGCATCAGCAGTGGCCGGGTCAITGG   | 1140                       |
| Qy                         | 1141  | CAAGGCTGAAAAACCGTGAACGAGTTGACGAGTTTGAACGAGTTCAGGTTGAGTACC    | 1200                       |
| Db                         | 1141  | CAAGGCTGAAAAACCGTGAACGAGTTGACGAGTTTGAACGAGTTCAGGTTGAGTACC    | 1200                       |
| Qy                         | 1201  | AAGAGCAGACCCCTGATGAGACGACAGGTCTCTGAAATCATCGACATTTCTA         | 1260                       |
| Db                         | 1201  | AAGAGCAGACCCCTGATGAGACGACAGGTCTCTGAAATCATCGACATTTCTA         | 1260                       |
| Qy                         | 1261  | TGCCAGTCAGATGGCTCAACGGAAGATCCGAGACATCTGCGCCAGGTTAAGCAGCAGCA  | 1320                       |
| Db                         | 1261  | TGCCAGTCAGATGGCTCAACGGAAGATCCGAGACATCTGCGCCAGGTTAAGCAGCAGCA  | 1320                       |
| Qy                         | 1321  | TCAGAAGGACAGAGTAAACAGGCGCCAGGCGAGGAGGAGTGAACAGCCCTTCCCTGTCC  | 1380                       |
| Db                         | 1321  | TCAGAAGGACAGAGTAAACAGGCGCCAGGCGAGGAGGAGTGAACAGCCCTTCCCTGTCC  | 1380                       |
| Qy                         | 1381  | CTTNGAGTCCAGGACAAACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG     | 1440                       |
| Db                         | 1381  | CTTNGAGTCCAGGACAAACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG     | 1440                       |
| Qy                         | 1441  | GAATGAGTGGGAAATCCGGGACACNTGGCGCGGCTGTAGATCAGGTTTGGCCACTTGATT | 1500                       |
| Db                         | 1441  | GAATGAGTGGGAAATCCGGGACACNTGGCGCGGCTGTAGATCAGGTTTGGCCACTTGATT | 1500                       |
| Qy                         | 1501  | GAGAAAGATGTTCCAGTGAAGAACCTGATCTNTCAGCCCAACACACCCAGGAGTGGC    | 1560                       |
| Db                         | 1501  | GAGAAAGATGTTCCAGTGAAGAACCTGATCTNTCAGCCCAACACACCCAGGAGTGGC    | 1560                       |
| Qy                         | 1561  | CCAAACACTGTTGCGGCTGTCAGAAATTTAGCGAAGGACATTTTAAACGTTGG        | 1620                       |
| Db                         | 1561  | CCAAACACTGTTGCGGCTGTCAGAAATTTAGCGAAGGACATTTTAAACGTTGG        | 1620                       |
| Qy                         | 1621  | ATTGTTTAAAGAGTCTCCAGGCGCCCAAGAGGGTGGATCACCTCAGTGGGAAGA       | 1680                       |
| Db                         | 1621  | ATTGTTTAAAGAGTCTCCAGGCGCCCAAGAGGGTGGATCACCTCAGTGGGAAGA       | 1680                       |
| Qy                         | 1681  | AAATTAATTTTCTTCAGGTTTAAAA                                    | 1708                       |
| Db                         | 1681  | AAATTAATTTTCTTCAGGTTTAAAA                                    | 1708                       |
| RESULT 2                   |   |  |                            |
| AR343073                   |   |  |                            |
| LOCUS                      | AR343073  | 1708 bp  | DNA linear PAT 17-AUG-2003 |
| DEFINITION                 | Sequence 5 from patent US 6576756.  |  |                            |
| ACCESSION                  | AR343073  |  |                            |
| VERSION                    | AR343073.1  | GI:33738475  |                            |
| KEYWORDS                   |   |  |                            |
| SOURCE                     | Unknown.  |  |                            |
| ORGANISM                   | Unknown.  |  |                            |
| REFERENCE                  | Unclassified.   |  |                            |
| AUTHORS                    | 1 (bases 1 to 1708)   |  |                            |
| TITLE                      | Chen, Y.-T.; Gure, A.; Tsang, S.; Stockert, E.; Jager, E.; Alexander, K. and Old, L.J.                  |  |                            |
| JOURNAL                    | Isolated nucleic acid molecule encoding cancer associated antigen, the antigen itself, and uses thereof |  |                            |
| FEATURES                   | Patent: US 6576756-A 5 10-JUN-2003;   |  |                            |
| source                     | Location/Qualifiers   |  |                            |
| 1. .1708                   |   |  |                            |

|   |     |             |           |             |              |              |                            |      |  |
|---|-----|-------------|-----------|-------------|--------------|--------------|----------------------------|------|--|
| /organism="unknown"<br>/mol_type="genomic DNA"                |     |             |           |             |              |              |                            |      |  |
| ORIGIN  |     |             |           |             |              |              |                            |      |  |
| Query Match   |     |             |           |             |              |              |                            |      |  |
| Best Local Similarity 99.7%; Score 1703; DB 6; Length 1708;   |     |             |           |             |              |              |                            |      |  |
| Matches 1708; Conservative 0; Mismatches 0; Indels 0; Gaps 0; |     |             |           |             |              |              |                            |      |  |
| Qy  | 1   | AGGAGCGTGC  | CGGACCGCC | CGCCAGTTT   | ACCCCGGGAGCC | CATCATGAAGCT | GAATGGCCA                  | 60   |  |
| Db  | 1   | AGGAGCGTGC  | CGGACCGCC | CGCCAGTTT   | ACCCCGGGAGCC | CATCATGAAGCT | GAATGGCCA                  | 60   |  |
| Qy  | 61  | CCAGTTGAGAA | CAACATGCT | CCCTGAAGTCT | CTCTACATCCC  | CGATGAGCAG   | TAGCAGAGG                  | 120  |  |
| Db  | 61  | CCAGTTGAGAA | CAACATGCT | CCCTGAAGTCT | CTCTACATCCC  | CGATGAGCAG   | TAGCAGAGG                  | 120  |  |
| Qy  | 121 | ACCTGAGAA   | TGGCGCGAG | GGGGCTTTGG  | CTCTCGGGGT   | CAGCCCGCC    | CAGGGCTCAC                 | 180  |  |
| Db  | 121 | ACCTGAGAA   | TGGCGCGAG | GGGGCTTTGG  | CTCTCGGGGT   | CAGCCCGCC    | CAGGGCTCAC                 | 180  |  |
| Qy  | 181 | TGTGGCAGC   | GGGGCCCC  | CAGCAAGCAG  | CAGCAAGTGA   | CATCCCCCT    | TGGCTCTGGT                 | 240  |  |
| Db  | 181 | TGTGGCAGC   | GGGGCCCC  | CAGCAAGCAG  | CAGCAAGTGA   | CATCCCCCT    | TGGCTCTGGT                 | 240  |  |
| Qy  | 241 | CCCCACCA    | GATGTGG   | GTGGTGCATT  | TATGGCAAG    | GGGGCCAC     | CATCCGCAACAT               | 300  |  |
| Db  | 241 | CCCCACCA    | GATGTGG   | GTGGTGCATT  | TATGGCAAG    | GGGGCCAC     | CATCCGCAACAT               | 300  |  |
| Qy  | 301 | AAAAACAG    | CCCCAGT   | CCCAAGAT    | AGTGCAT      | TAGGAAG      | GAGAACGAGT                 | 360  |  |
| Db  | 301 | AAAAACAG    | CCCCAGT   | CCCAAGAT    | AGTGCAT      | TAGGAAG      | GAGAACGAGT                 | 360  |  |
| Qy  | 361 | AGCCATCAG   | TGTGCAC   | TCCACCTC    | TGAGGGT      | GTCTCTCG     | CTTGTAAAGAT                | 420  |  |
| Db  | 361 | AGCCATCAG   | TGTGCAC   | TCCACCTC    | TGAGGGT      | GTCTCTCG     | CTTGTAAAGAT                | 420  |  |
| Qy  | 421 | GATTATGAT   | AAAGAGCT  | AAAGACAC    | CAAAACGG     | CTGACGAG     | TTCCCCCTGAAGAT             | 480  |  |
| Db  | 421 | GATTATGAT   | AAAGAGCT  | AAAGACAC    | CAAAACGG     | CTGACGAG     | TTCCCCCTGAAGAT             | 480  |  |
| Qy  | 481 | GGCCATTA    | ATAACTTT  | TGAGGGCT    | CTCATTTG     | GCAAGAG      | AGGACGTAAGAGGT             | 540  |  |
| Db  | 481 | GGCCATTA    | ATAACTTT  | TGAGGGCT    | CTCATTTG     | GCAAGAG      | AGGACGTAAGAGGT             | 540  |  |
| Qy  | 541 | AGAGCAAG    | ATACCGAG  | CAAAATCA    | CCATCTCT     | CTCTGTC      | CAAGACCTTACCCTTTACAA       | 600  |  |
| Db  | 541 | AGAGCAAG    | ATACCGAG  | CAAAATCA    | CCATCTCT     | CTCTGTC      | CAAGACCTTACCCTTTACAA       | 600  |  |
| Qy  | 601 | CCCTGAGAG   | CAATCAT   | CTGTGAAG    | GGGGCCAT     | CGAATTT      | GTTCGAGGGCCGAGCAGGA        | 660  |  |
| Db  | 601 | CCCTGAGAG   | CAATCAT   | CTGTGAAG    | GGGGCCAT     | CGAATTT      | GTTCGAGGGCCGAGCAGGA        | 660  |  |
| Qy  | 661 | AATAATGA    | AGAAAGT   | TCGGAGG     | CGCTAT       | GAGATGAT     | GTGGTGGCCATGAGCTCACCT      | 720  |  |
| Db  | 661 | AATAATGA    | AGAAAGT   | TCGGAGG     | CGCTAT       | GAGATGAT     | GTGGTGGCCATGAGCTCACCT      | 720  |  |
| Qy  | 721 | GATCCCTG    | GGCTGA    | ACTGCTGT    | GTAGTCTTT    | TTCCAGCT     | TCATCCAGCGCAGTCCC          | 780  |  |
| Db  | 721 | GATCCCTG    | GGCTGA    | ACTGCTGT    | GTAGTCTTT    | TTCCAGCT     | TCATCCAGCGCAGTCCC          | 780  |  |
| Qy  | 781 | GCGCCCTC    | CCAGCAG   | CGCTTACT    | GGGCTGCT     | CCCTATAG     | CTCTTTATGAGGCTCCCGA        | 840  |  |
| Db  | 781 | GCGCCCTC    | CCAGCAG   | CGCTTACT    | GGGCTGCT     | CCCTATAG     | CTCTTTATGAGGCTCCCGA        | 840  |  |
| Qy  | 841 | GCAGGAGT    | GTGAGT    | GTGTTAT     | TCGCCGCC     | CAGGCGAG     | TGGGGCCATCATCGGCAAGAA      | 900  |  |
| Db  | 841 | GCAGGAGT    | GTGAGT    | GTGTTAT     | TCGCCGCC     | CAGGCGAG     | TGGGGCCATCATCGGCAAGAA      | 900  |  |
| Qy  | 901 | GGGGCAGC    | ACATCAA   | CAGCTCT     | CCCGTTT      | TGCGAG       | CGCTCCATCAAGATTGCAAC       | 960  |  |
| Db  | 901 | GGGGCAGC    | ACATCAA   | CAGCTCT     | CCCGTTT      | TGCGAG       | CGCTCCATCAAGATTGCAAC       | 960  |  |
| Qy  | 961 | CGAAACAC    | CTGACT    | CCAAAGT     | TCGTATG      | TGTTAT       | CATCACTGGAGCCGCGAGGCCCAATT | 1020 |  |

BD209924 Isolated nucleic acid molecules encoding cancer-associated antigens, these antigens and method of using the same.  
BD209924  
BD209924.1 GI:33019694  
JP 2002512049-A/3  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Chen, Y.T., Gure, A., Tsang, S., Stockert, E., Jager, E., Knuth, A. and Old, L.J.  
Isolated nucleic acid molecules encoding cancer-associated antigens, these antigens and method of using the same  
Patent: JP 2002512049-A 3 23-APR-2002;  
LUDWIG INSTITUTE FOR CANCER RESEARCH  
OS Homo sapiens (human)  
PN JP 2002512049-A/3

RESULT 3  
BD209924  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

|    |      |          |         |         |         |        |          |         |               |      |
|----|------|----------|---------|---------|---------|--------|----------|---------|---------------|------|
| Db | 961  | CGAAACAC | CTGACT  | CCAAAGT | TCGTATG | TGTTAT | CATCACT  | GGAGCCG | CAGAGGCCCAATT | 1020 |
| Qy | 1021 | CRAGGCTC | AGGAGAG | ATCTAT  | TGGCAAA | CTCAAG | GAGGAACT | TTCTTTT | GGTCCCAGGA    | 1080 |
| Db | 1021 | CRAGGCTC | AGGAGAG | ATCTAT  | TGGCAAA | CTCAAG | GAGGAACT | TTCTTTT | GGTCCCAGGA    | 1080 |
| Qy | 1081 | GGAAGTGA | AGCTGG  | AGACCC  | ACATAG  | CTGTGC | CAGCAT   | CAGCAG  | CTGGCGGGTCA   | 1140 |
| Db | 1081 | GGAAGTGA | AGCTGG  | AGACCC  | ACATAG  | CTGTGC | CAGCAT   | CAGCAG  | CTGGCGGGTCA   | 1140 |
| Qy | 1141 | CAAAGTGA | AAAAA   | CGGTGAA | CGAGTT  | GCAGAT | TTGAC    | CGGAGT  | GGTGGTATG     | 1200 |
| Db | 1141 | CAAAGTGA | AAAAA   | CGGTGAA | CGAGTT  | GCAGAT | TTGAC    | CGGAGT  | GGTGGTATG     | 1200 |
| Qy | 1201 | AAGAGACA | GAGAC   | CCCTGT  | ATGAGAA | CGAC   | CAGGTC   | ATCGT   | GAATTTCTA     | 1260 |
| Db | 1201 | AAGAGACA | GAGAC   | CCCTGT  | ATGAGAA | CGAC   | CAGGTC   | ATCGT   | GAATTTCTA     | 1260 |
| Qy | 1261 | TGCCAGT  | CAAGT   | GGTCAA  | CGGAAG  | ATCCG  | AGACAT   | TCCTG   | CCCGAGT       | 1320 |
| Db | 1261 | TGCCAGT  | CAAGT   | GGTCAA  | CGGAAG  | ATCCG  | AGACAT   | TCCTG   | CCCGAGT       | 1320 |
| Qy | 1321 | TCAGAA   | GGGAC   | AGAGTA  | ACCAGG  | CCCG   | CAGGAG   | TGAC    | CAGCCCTC      | 1380 |
| Db | 1321 | TCAGAA   | GGGAC   | AGAGTA  | ACCAGG  | CCCG   | CAGGAG   | TGAC    | CAGCCCTC      | 1380 |
| Qy | 1381 | CTTNGAG  | TCCAG   | CAACAA  | CGGGC   | AGAA   | ATCG     | AGAT    | GTGCTC        | 1440 |
| Db | 1381 | CTTNGAG  | TCCAG   | CAACAA  | CGGGC   | AGAA   | ATCG     | AGAT    | GTGCTC        | 1440 |
| Qy | 1441 | GAATGAG  | TGGGA   | ATCCG   | GACAC   | NTG    | GGCCGG   | CTGT    | AGATCAG       | 1500 |
| Db | 1441 | GAATGAG  | TGGGA   | ATCCG   | GACAC   | NTG    | GGCCGG   | CTGT    | AGATCAG       | 1500 |
| Qy | 1501 | GAGAA    | GATG    | TTTCC   | AGTGAG  | GAAC   | CCCTG    | ATCT    | NTCAG         | 1560 |
| Db | 1501 | GAGAA    | GATG    | TTTCC   | AGTGAG  | GAAC   | CCCTG    | ATCT    | NTCAG         | 1560 |
| Qy | 1561 | CCAA     | CATG    | TNTG    | CCCTC   | CGGGT  | GT       | CAGAA   | TTTAG         | 1620 |
| Db | 1561 | CCAA     | CATG    | TNTG    | CCCTC   | CGGGT  | GT       | CAGAA   | TTTAG         | 1620 |
| Qy | 1621 | ATTG     | TTTAA   | AGAG    | CTCT    | CCAG   | GGCC     | CCAC    | CAAG          | 1680 |
| Db | 1621 | ATTG     | TTTAA   | AGAG    | CTCT    | CCAG   | GGCC     | CCAC    | CAAG          | 1680 |
| Qy | 1681 | AAAA     | TAAAT   | TTTCT   | TCAG    | GGTTT  | TAAAA    | 1708    |               |      |
| Db | 1681 | AAAA     | TAAAT   | TTTCT   | TCAG    | GGTTT  | TAAAA    | 1708    |               |      |

```

PD      23-APR-2002
PF      16-MAR-1999 JP 2000545030
PR      17-APR-1998 US 09/061709
PI      YAO TSENG CHEN,ALI GURE,SOLAM TSANG,ELISABETH STOCKERT,ELKE
PI      JAGER,
PI      ALEXANDER KNUTH,LLOYD J OLD
PC      C12N15/09,A61K35/12,A61K39/00,A61K39/39,A61P35/00,C07K16/32,
PC      C12N1/15,
PC      C12N1/19,C12N1/21,C12N5/10,C12P21/08,C12Q1/68,G01N33/53,G01N33/ PC
574,
PC      C12N15/00,C12N5/00
CC      n=not determined(1384,1464,1533,1571,1595)
FH      Key Location/Qualifiers
FT      source 1..1708
FT      Location/Qualifiers
FT      /organism='Homo sapiens (human)'.
FT      source 1..1708
FT      /organism='Homo sapiens'
FT      /mol_type='genomic DNA'
FT      /db_xref='taxon:9606'

ORIGIN
Query Match 99.7%; Score 1703; DB 6; Length 1708;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AGGCAGCGTGGCGCACCGCCCGCTTTTACCCCGGGAGCCATCATGAAGCTGAATGGCCA 60
DB      1 AGGCAGCGTGGCGCACCGCCCGCTTTTACCCCGGGAGCCATCATGAAGCTGAATGGCCA 60

QY      61 CCAGTTGGAGAACCATGCCCTGAAGTCTCTACATCCCGATGAGCAGATACACAGGG 120
DB      61 CCAGTTGGAGAACCATGCCCTGAAGTCTCTACATCCCGATGAGCAGATACACAGGG 120

QY      121 ACCTGAGAAATGGCGCGAGGGGCTTTGGCTCTCGGGGTGAGCCCGCCAGGGCTCACC 180
DB      121 ACCTGAGAAATGGCGCGAGGGGCTTTGGCTCTCGGGGTGAGCCCGCCAGGGCTCACC 180

QY      181 TGTGGCAGCGGGGCGCCCGCAGCAGAGCAGAGTGGACATCCCGCTTCGGCTCCTGGT 240
DB      181 TGTGGCAGCGGGGCGCCCGCAGCAGAGCAGAGTGGACATCCCGCTTCGGCTCCTGGT 240

QY      241 GCCCACCAGTATGTGGTGGCTATTATTTGGCAAGGAGGGGCCACCATCCGCAACATCAC 300
DB      241 GCCCACCAGTATGTGGTGGCTATTATTTGGCAAGGAGGGGGCCACCATCCGCAACATCAC 300

QY      301 AAAACAGACCCAGTCCAAAGATAGACGTGCATAGGAAGGAGAACGCGAGTGCGAGTGA 360
DB      301 AAAACAGACCCAGTCCAAAGATAGACGTGCATAGGAAGGAGAACGCGAGTGCGAGTGA 360

QY      361 AGCCATCAGTGTGACATCCACCCCTGAGGGTGTCTCTCGCTTTGAAGATGATCTTGA 420
DB      361 AGCCATCAGTGTGACATCCACCCCTGAGGGTGTCTCTCGCTTTGAAGATGATCTTGA 420

QY      421 GATTATGATAAAGAGGTGAAGGACACAAACCGGTGACGAGGTTCCTCGTGAAGATCCT 480
DB      421 GATTATGATAAAGAGGTGAAGGACACAAACCGGTGACGAGGTTCCTCGTGAAGATCCT 480

QY      481 GGCCCATTAATCTTTGTAGGGCTCTCATTTGGCAAGGAGGAGCGAACCTGAGAGGT 540
DB      481 GGCCCATTAATCTTTGTAGGGCTCTCATTTGGCAAGGAGGAGCGAACCTGAGAGGT 540

QY      541 AGAGCAAGATACCGAGACAAAATACCATCTCTCGTTGCAAGACCTTACCCTTTACAA 600
DB      541 AGAGCAAGATACCGAGACAAAATACCATCTCTCGTTGCAAGACCTTACCCTTTACAA 600

QY      601 CCCTGAGAGGACCATCACTGTGAAGGGGCCCATCGAATTTGTGAGGGCCGAGCAGGA 660
DB      601 CCCTGAGAGGACCATCACTGTGAAGGGGGCCCATCGAATTTGTGAGGGCCGAGCAGGA 660

QY      661 AATAATGAAGAAAGTTCCGGAGGGCTATGAGATGATGTGGCTGCCATGAGCTCTCACCT 720

```

LOCUS AR171866 1946 bp DNA linear PAT 17-DEC-2001  
DEFINITION Sequence 7 from patent US 6297364.  
ACCESSION AR171866  
VERSION AR171866.1 GI:17910816  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1. (bases 1 to 1946)  
Chen, Y.-T., Gure, A., Tsang, S., Stockert, E., Jager, E., Alexander, K.  
and Old, L. J.  
TITLE Isolated nucleic acid molecule encoding cancer associated antigen,  
the antigen itself, and uses thereof  
JOURNAL Patent: US 6297364-A 7 02-OCR-2001;  
FEATURES Location/Qualifiers  
source 1..1946  
/organism="unknown"  
/mol\_type="unassigned DNA"  
ORIGIN  
Query Match 97.7%; Score 1669.2; DB 6; Length 1946;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1676; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 30 CCCGGGAGCCATCATGAAGCTGAATGGCCACCAAGTTGGAGAACCATGCGCTCGAAGGTCT 89  
DB 268 CCAGGCAAGCCATCATGAAGCTGAATGGCCACCAAGTTGGAGAACCATGCGCTCGAAGGTCT 327  
QY 90 CCTACATCCCGATGAGCAGATACACAGGACCTGAGATGGCGCGGAGGGGCTTTG 149  
DB 328 CCTACATCCCGGATGAGCAGATACACAGGACCTGAGATGGCGCGGAGGGGCTTTG 387  
QY 150 GCTCTCGGGGTACGCCCCGCGCAGGGCTCACCTGTGGCAGGGGGGCCCCAGCAAGCAGC 209  
DB 388 GCTCTCGGGGTACGCCCCGCGCAGGGCTCACCTGTGGCAGGGGGGCCCCAGCAAGCAGC 447  
QY 210 AGCAAGTGGACATCCCCCTTCGGTCTCGTGTGCGCCACCCAGTATGTGGGTGCCATTATTC 269  
DB 448 AGCAAGTGGACATCCCCCTTCGGTCTCGTGTGCGCCACCCAGTATGTGGGTGCCATTATTC 507  
QY 270 GCAGGAGGGGGCCACCATCCGCAACATCACAAACAGCCAGTGTGCATCCACCCCTGAGG 329  
DB 508 GCAGGAGGGGGCCACCATCCGCAACATCACAAACAGCCAGTGTGCATCCACCCCTGAGG 567  
QY 330 ATAGGAAGGAGAACGCGAGTGCAGCTGAAAGAGCCATCAGTGTGCATCCACCCCTGAGG 389  
DB 568 ATAGGAAGGAGAACGCGAGTGCAGCTGAAAGAGCCATCAGTGTGCATCCACCCCTGAGG 627  
QY 390 GCTGCTCTCGCTGTGAAGATGATCTTGGAGATATGCAATAGAGGCTTAAGCACCA 449  
DB 628 GCTGCTCTCGCTGTGAAGATGATCTTGGAGATATGCAATAGAGGCTTAAGCACCA 687  
QY 450 AAACGGCTGACGAGGTTCCCTGAAGATCTTGGCCCATATAACTTTGTAGGGGCTCTCA 509  
DB 688 AAAAGGCTGACGAGGTTCCCTGAAGATCTTGGCCCATATAACTTTGTAGGGGCTCTCA 747  
QY 510 TTGCAAGAGGAGCGAAACCTGAAGAGGTGAGAGCAAGATACCGAGACAAAATACCA 569  
DB 748 TTGCAAGAGGAGCGAAACCTGAAGAGGTGAGAGCAAGATACCGAGACAAAATACCA 807  
QY 570 TCTCTCTGTGCAAGACTTACCTTTACACCTGAGAGGACCATCACCTGTGAAGGGG 629  
DB 808 TCTCTCTGTGCAAGACTTACCTTTACACCTGAGAGGACCATCACCTGTGAAGGGG 867  
QY 630 CCATCGAAGATTGTGAGGGCCGAGCAGGAAATAATGAAGAAAGTTTCGGGAGGCGCTATG 689  
DB 868 CCATCGAAGATTGTGAGGGCCGAGCAGGAAATAATGAAGAAAGTTTCGGGAGGCGCTATG 927  
QY 690 AGAATGATGGCTGCCATGAGCTCTACCTGATCCCTGGCTGAACTGGCTGCTGTAG 749  
DB 928 AGAATGATGGCTGCCATGAGCTCTACCTGATCCCTGGCTGAACTGGCTGCTGTAG 987  
QY 750 GTCCTTTCCAGCTTCATCCAGCGCAGTCCCGCGCGCTCCAGCAGCGGTTACTGGGGCTG 809

DB 988 GTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCCCGCTCCCGAGCAGCGTTACTGGGGTGG 1047  
QY 810 CTCCTCTATAGCTCCTTTATGTCAGGCTCCCGACGAGAGATGCTGAGGTGTTTATCCCGG 869  
DB 1048 CTCCTCTATAGCTCCTTTATGTCAGGCTCCCGACGAGAGATGCTGAGGTGTTTATCCCG 1107  
QY 870 CCCAGGCACTGGGCGCATCATCGGCAAGAGAGGGGAGCACAATCAAAAGCTCTCCCGGT 929  
DB 1108 CCCAGGCACTGGGCGCATCATCGGCAAGAGAGGGGAGCACAATCAAAAGCTCTCCCGGT 1167  
QY 930 TTGCCAGCGCTCCATCAAGATTGACACCCGAAACACCTGACTCCAAAGTTCTGATGG 989  
DB 1168 TTGCCAGCGCTCCATCAAGATTGACACCCGAAACACCTGACTCCAAAGTTCTGATGG 1227  
QY 990 TTATCATCTGAGCGCGCAGAGGCCCAATTCAAAGCTCAGGGAAGAAATCTATGCAAAAC 1049  
DB 1228 TTATCATCTGAGCGCGCAGAGGCCCAATTCAAAGCTCAGGGAAGAAATCTATGCAAAAC 1287  
QY 1050 TCAAGAGGAGAACTCTTTGCTCCNAGGAGGAGTGAAGCTGAGACCCCAATACCTG 1109  
DB 1288 TCAAGAGGAGAACTCTTTGCTCCNAGGAGGAGTGAAGCTGAGACCCCAATACCTG 1347  
QY 1110 TGCAGCATCAGCAGCTGCGCGGTCAITGGCAAAAGGTGGAATAACGGTGAACGAGTTGC 1169  
DB 1348 TGCAGCATCAGCAGCTGCGCGGTCAITGGCAAAAGGTGGAATAACGGTGAACGAGTTGC 1407  
QY 1170 AGAATTTGACGGCAGCTGAGGTGAGTACCAAGAGACAGACCCCTGATGAGAACGACC 1229  
DB 1408 AGAATTTGACGGCAGCTGAGGTGAGTACCAAGAGACAGACCCCTGATGAGAACGACC 1467  
QY 1230 AGGTCTATCTGAAATCATCGGACATTTCTATGCCAGTCAGATGGTCAACCGAGAGATCC 1289  
DB 1468 AGGTCTATCTGAAATCATCGGACATTTCTATGCCAGTCAGATGGTCAACCGAGAGATCC 1527  
QY 1290 GAGACATCTCGCCAGGTTAAGCAGCAGCATCAGAGGGAAGAGAGGAGGAGGAGGAGG 1349  
DB 1528 GAGACATCTCGCCAGGTTAAGCAGCAGCATCAGAGGGAAGAGAGGAGGAGGAGGAGG 1587  
QY 1350 CACGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1409  
DB 1588 CACGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1647  
QY 1410 ATCGAGAGTGTGCTCTCCCGGCGAGGCTGAGAAATGAGTGGGAATCCCGGAGACACNTGGG 1469  
DB 1648 ATCGAGAGTGTGCTCTCCCGGCGAGGCTGAGAAATGAGTGGGAATCCCGGAGACACNTGGG 1707  
QY 1470 CGGGCTGTAGATCAGGTTTGGCCACTTGTGATGAGAAAGATGTTCCAGTGAGGAAACCTGA 1529  
DB 1708 CGGGCTGTAGATCAGGTTTGGCCACTTGTGATGAGAAAGATGTTCCAGTGAGGAAACCTGA 1767  
QY 1530 TCTNTCAGCCCCCAACACACCCCAATGGCCCAACACTGTTGCCCCCTCGGGGTGTGAG 1589  
DB 1768 TCTNTCAGCCCCCAACACACCCCAATGGCCCAACACTGTTGCCCCCTCGGGGTGTGAG 1827  
QY 1590 AAATNTAGCGAAGGCACTTTTAAACGTTGATTTTAAAGAAAGCTCTCCAGGCCCCAC 1649  
DB 1828 AAATNTAGCGAAGGCACTTTTAAACGTTGATTTTAAAGAAAGCTCTCCAGGCCCCAC 1887  
QY 1650 CAAGAGGGTGTATCACCTCAGTGGGAAGAAAATAAAATTTCTTCAGGTTTAAAAA 1708  
DB 1888 CAAGAGGGTGTATCACCTCAGTGGGAAGAAAATAAAATTTCTTCAGGTTTAAAAA 1946

## RESULT 5

LOCUS AR343075 1946 bp DNA linear PAT 17-AUG-2003  
DEFINITION Sequence 7 from patent US 6576756.  
ACCESSION AR343075  
VERSION AR343075.1 GI:33738477  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.

Unclassified.  
 1 (bases 1 to 1946)  
 Chen, Y.-T., Gure, A., Tsang, S., Stockert, E., Jager, E., Alexander, K.  
 and Old, L.J.  
 Isolated nucleic acid molecule encoding cancer associated antigen,  
 the antigen itself, and uses thereof  
 Patent: US 6576756-A 7 10-JUN-2003;  
 JOURNAL  
 FEATURES  
 Location/Qualifiers  
 source  
 1..1946  
 /organism="unknown"  
 /mol\_type="genomic DNA"

ORIGIN

Query Match 97.7%; Score 1669.2; DB 6; Length 1946;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 1676; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

30 CCGGGAGCCATCATGAAGCTGAATGGCCACCACTGGAGAACCATGCCCTGAAGTCT 89  
 Db |||||  
 268 CCAGGCAAGCCATCATGAAGCTGAATGGCCACCACTGGAGAACCATGCCCTGAAGTCT 327  
 Qy |||||  
 90 CCTACATCCCCGATCAGCAGATAGCACAGGACCTGAGATGGCGCCGAGGGGGCTTTG 149  
 Db |||||  
 328 CCTACATCCCCGATCAGCAGATAGCACAGGACCTGAGATGGCGCCGAGGGGGCTTTG 387  
 Qy |||||  
 150 GCTCTCGGGGTACGCCCGCCAGGGCTCACCTGTGGCAGCGGGGGCCCAAGCAGC 209  
 Db |||||  
 388 GCTCTCGGGGTACGCCCGCCAGGGCTCACCTGTGGCAGCGGGGGCCCAAGCAGC 447  
 Qy |||||  
 210 AGCAAGTGGACATCCCTCTCGGCTCTGTGGTCCACCCAGTATGTGGGTGCCATTATTG 269  
 Db |||||  
 448 AGCAAGTGGACATCCCTCTCGGCTCTGTGGTCCACCCAGTATGTGGGTGCCATTATTG 507  
 Qy |||||  
 270 GCAAGGGGGGGCCACCATCGCAACATCAAAACAGACCCAGTCCCAAGATAGACGTGC 329  
 Db |||||  
 508 GCAAGGGGGGGCCACCATCGCAACATCAAAACAGACCCAGTCCCAAGATAGACGTGC 567  
 Qy |||||  
 330 ATAGAGAGGAGACGCGAGGTGCAGCTGAAAAGCCATCAGTGTGCATCCACCCCTGAG 389  
 Db |||||  
 568 ATAGAGAGGAGACGCGAGGTGCAGCTGAAAAGCCATCAGTGTGCATCCACCCCTGAG 627  
 Qy |||||  
 390 GCTGCTCCTCGCTGTGAAGATGATCTTGGAGATTATGCAATAAGAGGCTTAAGCACCA 449  
 Db |||||  
 628 GCTGCTCCTCGCTGTGAAGATGATCTTGGAGATTATGCAATAAGAGGCTTAAGCACCA 687  
 Qy |||||  
 450 AAAGCGGTGACGAGGTTCCTCGAAGATCTCGCCCATTAATTAATCTTGTAGGGCGTCTCA 509  
 Db |||||  
 688 AAAGCGGTGACGAGGTTCCTCGAAGATCTCGCCCATTAATTAATCTTGTAGGGCGTCTCA 747  
 Qy |||||  
 510 TTGCAAGAGGAGGACGAACTGAAGAGGTAGAGCAAGATACCGAGACAAAATCACCA 569  
 Db |||||  
 748 TTGCAAGAGGAGGACGAACTGAAGAGGTAGAGCAAGATACCGAGACAAAATCACCA 807  
 Qy |||||  
 570 TCTCCTCGTTCARAGACCTTACCCCTTTACAACTCTGAGAGGACCATCACTGTGAAGGGG 629  
 Db |||||  
 808 TCTCCTCGTTCARAGACCTTACCCCTTTACAACTCTGAGAGGACCATCACTGTGAAGGGG 867  
 Qy |||||  
 630 CCATCGAGAAATTGTTGAGGGCCGAGCAGGAAATATGAAGAAAGTTCCGGAGGCGCTATG 689  
 Db |||||  
 868 CCATCGAGAAATTGTTGAGGGCCGAGCAGGAAATATGAAGAAAGTTCCGGAGGCGCTATG 927  
 Qy |||||  
 690 AGAATGATGTGGTGCATGAGCTCTCACTGTATCCCTGCTGAACTGGCTGCTGTAG 749  
 Db |||||  
 928 AGAATGATGTGGTGCATGAGCTCTCACTGTATCCCTGCTGAACTGGCTGCTGTAG 987  
 Qy |||||  
 750 GTCTTTTCCAGCTTCATCAGCGCAGTCCGGCGCTCCAGCAGGTTACTGGGGCTG 809  
 Db |||||  
 988 GTCTTTTCCAGCTTCATCAGCGCAGTCCGGCGCTCCAGCAGGTTACTGGGGCTG 1047  
 Qy |||||  
 810 CTCCTCTAGTCTCTTTATGAGGCTCCCGAGCAGGAGATGGTGCAGGTGTTTATCCCCG 869  
 Db |||||  
 1048 CTCCTCTAGTCTCTTTATGAGGCTCCCGAGCAGGAGATGGTGCAGGTGTTTATCCCCG 1107

870 CCCAGGCAGTGGCGCCCATCATCGCAAGAAAGGGGCGACATCAAAACAGCTCTCCCGGT 929  
 Db |||||  
 1108 CCCAGGCAGTGGCGCCCATCATCGCAAGAAAGGGGCGACATCAAAACAGCTCTCCCGGT 1167  
 Qy |||||  
 930 TTGCCAGGCGCTCCATCAAGATTTGACACCCCGAAACACCTGACTCCAAAGTTGGTATGG 989  
 Db |||||  
 1168 TTGCCAGGCGCTCCATCAAGATTTGACACCCCGAAACACCTGACTCCAAAGTTGGTATGG 1227  
 Qy |||||  
 990 TTATCATCACTGGACCGCCAGAGGCCCAATTTCAAGGCTCAGGGAGAAATCTATGGCAAA 1049  
 Db |||||  
 1228 TTATCATCACTGGACCGCCAGAGGCCCAATTTCAAGGCTCAGGGAGAAATCTATGGCAAA 1287  
 Qy |||||  
 1050 TCAAGGAGGAGAACTTTCTTGGTCCCAGAGGAGTGAAGCTCGAGACCCACATAGCTG 1109  
 Db |||||  
 1288 TCAAGGAGGAGAACTTTCTTGGTCCCAGAGGAGTGAAGCTCGAGACCCACATAGCTG 1347  
 Qy |||||  
 1110 TGCCAGCATCAGCAGCTGGCGGGTCAATTGGCAAGGTGGAAACACGTTGAACGAGTTGC 1169  
 Db |||||  
 1348 TGCCAGCATCAGCAGCTGGCGGGTCAATTGGCAAGGTGGAAACACGTTGAACGAGTTGC 1407  
 Qy |||||  
 1170 AGAATTTGACGCGCAGCTGAGGTGTAGTACCAAGAGACCAAGCCCTGATGAGAACGACC 1229  
 Db |||||  
 1408 AGAATTTGACGCGCAGCTGAGGTGTAGTACCAAGAGACCAAGCCCTGATGAGAACGACC 1467  
 Qy |||||  
 1230 AGGTCACTCGTGAATCATCGGACATTTCTATGTCAGTCAGATGGCTCAACGGAGATCC 1289  
 Db |||||  
 1468 AGGTCACTCGTGAATCATCGGACATTTCTATGTCAGTCAGATGGCTCAACGGAGATCC 1527  
 Qy |||||  
 1290 GAGACATCTCGGCCCAGGTTAAGCAGCAGCATCAAGAGGAGCAGAGTAAACAGGCCCCAGG 1349  
 Db |||||  
 1528 GAGACATCTCGGCCCAGGTTAAGCAGCAGCATCAAGAGGAGCAGAGTAAACAGGCCCCAGG 1587  
 Qy |||||  
 1350 CACGAGGAGTGTACAGCCCTCCTGTCTCCCTTNGAGTCAGGACCAACAGCGGCGAGAA 1409  
 Db |||||  
 1588 CACGAGGAGTGTACAGCCCTCCTGTCTCCCTTNGAGTCAGGACCAACAGCGGCGAGAA 1647  
 Qy |||||  
 1410 ATCGAGGTGTCTCTCCCGCAGGCTGAGATGAGTGGGAATCGGGACACNTGGGC 1469  
 Db |||||  
 1648 ATCGAGGTGTCTCTCCCGCAGGCTGAGATGAGTGGGAATCGGGACACNTGGGC 1707  
 Qy |||||  
 1470 CGGCTGTAGATCAGGTTTGCCACTTGATTTAGAAAGATGTTCCAGTAGGAAACCTCTGA 1529  
 Db |||||  
 1708 CGGCTGTAGATCAGGTTTGCCACTTGATTTAGAAAGATGTTCCAGTAGGAAACCTCTGA 1767  
 Qy |||||  
 1530 TCTNTCAGCCCCAACACCCCAATTTGGCCCACTGNTGCCCTCGGGGTGTGAG 1589  
 Db |||||  
 1768 TCTNTCAGCCCCAACACCCCAATTTGGCCCACTGNTGCCCTCGGGGTGTGAG 1827  
 Qy |||||  
 1590 AAATNTAGCGCAAGGCACCTTTTAAACGTGATTTTAAAGAAAGCTCTCCAGGCCCCAC 1649  
 Db |||||  
 1828 AAATNTAGCGCAAGGCACCTTTTAAACGTGATTTTAAAGAAAGCTCTCCAGGCCCCAC 1887  
 Qy |||||  
 1650 CAAGAGGTGTGATCAGCTCAGTGGGAGAAATAAATTTCTTCAGGTTTAAAA 1708  
 Db |||||  
 1888 CAAGAGGTGTGATCAGCTCAGTGGGAGAAATAAATTTCTTCAGGTTTAAAA 1946

RESULT 6  
 LOCUS BD209926  
 DEFINITION Isolated nucleic acid molecules encoding cancer-associated antigens, these antigens and method of using the same.  
 ACCESSION BD209926  
 VERSION BD209926.1 GI:33019696  
 KEYWORDS JP 2002512049-A/5.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1946)  
 AUTHORS Chen, Y.T., Gure, A., Tsang, S., Stockert, E., Jager, E., Knuth, A. and Old, L.J.  
 TITLE Isolated nucleic acid molecules encoding cancer-associated



|          |  |   |     |   |      |  |  |  |  |  |
|----------|--|---|-----|---|------|--|--|--|--|--|
| JOURNAL  | antigens, these antigens and method of using the same  |   |     |   |      |  |  |  |  |  |
| COMMENT  | Patent: JP 2002512049-A 5 23-APR-2002;<br>LUDWIG INSTITUTE FOR CANCER RESEARCH<br>OS Homo sapiens (human)<br>PN JP 2002512049-A/5<br>PD 23-APR-2002<br>PF 16-MAR-1999 JP 2000545030<br>PR 17-APR-1998 US 09/061709<br>PI YAO TSENG CHEN, ALI GURE, SOLAM TSANG, ELISABETH STOCKERT, ELKE JAGER,<br>PI ALEXANDER KNUTH, LLOYD J OLD<br>PC C12N15/09, A61K35/12, A61K39/00, A61K39/39, A61P35/00, C07K16/32, C12N1/15,<br>PC C12N1/19, C12N1/21, C12N5/10, C12P21/08, C12Q1/68, G01N33/53, G01N33/ PC 574,<br>PC C12N15/00, C12N5/00<br>CC n-not determined(1622,1702,1771,1809,1833)<br>FH Key Location/Qualifiers<br>FT source 1. .1946<br>FT Location/Qualifiers<br>1. .1946<br>/organism="Homo sapiens (human)".<br>/organism="Homo sapiens"<br>/mol_type="genomic DNA"<br>/db_xref="taxon:9606" |   |     |   |      |  |  |  |  |  |
| FEATURES | source   |   |     |   |      |  |  |  |  |  |
| ORIGIN   | Query Match 97.7%; Score 1669.2; DB 6; Length 1946;<br>Best Local Similarity 99.8%; Pred. No. 0;<br>Matches 1676; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  |   |     |   |      |  |  |  |  |  |
| QY       | 30   | CCCCGGAGCCATCATGAAGCTGAATGGCCACCAGTTGGGAACCATGCCCTGAAGTCT     | 89  | CCATCGAGAAATGTTGCGAGGCGGAGAGGAAATATGAAGAAAGTTGGGAGGCGTATG     | 927  |  |  |  |  |  |
| DB       | 268  | CCAGCGAAGCCATCATGAGCTGAATGGCCACCAGTTGGGAACCATGCCCTGAAGTCT     | 327 | AGAAATGATCTGGCTGCCATGAGCTCTCACCTGATCCCTGGCCTGAACCTGGCTGTGTAG  | 749  |  |  |  |  |  |
| QY       | 90   | CCTACATCCCCGATGAGCAGATAGCAGAGGACCTGAGAAATGGCGCGGAGGGGCTTG     | 149 | AGATGATGATGCTGCTGCCATGAGCTCTCACCTGATCCCTGGCCTGAACCTGGCTGTGTAG | 987  |  |  |  |  |  |
| DB       | 328  | CCTACATCCCCGATGAGCAGATAGCAGAGGACCTGAGAAATGGCGCGGAGGGGCTTG     | 387 | GTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCCGCTCCCAAGCAGCGTCTACTGGGGCTG  | 809  |  |  |  |  |  |
| QY       | 150  | GCTCTCGGGTCAGCCCGCAGAGGCTCACCTGTGGCAGCGGGGCCCCAGCAGC          | 209 | GTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCCGCTCCCAAGCAGCGTCTACTGGGGCTG  | 1047 |  |  |  |  |  |
| DB       | 388  | GCTCTCGGGTCAGCCCGCAGAGGCTCACCTGTGGCAGCGGGGCCCCAGCAGC          | 447 | CTCCCTATAGCTCCTTTATGTCAGGCTCCGAGCAGAGATGCTGCAGGTTGTTATCCCGG   | 869  |  |  |  |  |  |
| QY       | 210  | AGCAAGTGGACATCCCCCTCTCGGCTCCTGTGGTCCCAACCCAGTATGTGGTGCCTATTG  | 269 | CTCCCTATAGCTCCTTTATGTCAGGCTCCGAGCAGAGATGCTGCAGGTTGTTATCCCGG   | 1107 |  |  |  |  |  |
| DB       | 448  | AGCAAGTGGACATCCCCCTCTCGGCTCCTGTGGTCCCAACCCAGTATGTGGTGCCTATTG  | 507 | CCAGGCACTGCGGCCCATTCGCGAGCTGATGACCAAGAGACAGACCTGATGAGAACGCC   | 929  |  |  |  |  |  |
| QY       | 270  | GCAAGGAGGGGCGACATCCGCAACATCAAAACAGACCCAGTCCCAAGATAGACGTGC     | 329 | CCAGGCACTGCGGCCCATTCGCGAGCTGATGACCAAGAGACAGACCTGATGAGAACGCC   | 1229 |  |  |  |  |  |
| DB       | 508  | GCAAGGAGGGGCGACATCCGCAACATCAAAACAGACCCAGTCCCAAGATAGACGTGC     | 567 | CCAGGCACTGCGGCCCATTCGCGAGCTGATGACCAAGAGACAGACCTGATGAGAACGCC   | 1467 |  |  |  |  |  |
| QY       | 330  | ATAGGAAGGAGAACGACAGTGCAGTGAAGGACCATCAGTGTGCATCCACCCCTCAGG     | 389 | AGATGATCTCTGAAAATCATCGGACATTTCTATGCCAGTCAGATGGTCAACCGAAGATCC  | 1289 |  |  |  |  |  |
| DB       | 568  | ATAGGAAGGAGAACGACAGTGCAGTGAAGGACCATCAGTGTGCATCCACCCCTCAGG     | 627 | AGGTCACTGTGAAAATCATCGGACATTTCTATGCCAGTCAGATGGTCAACCGAAGATCC   | 1527 |  |  |  |  |  |
| QY       | 390  | GCTGCTCTCGCTGTGAAGATGATCTTGGAGATTATGATATAAGAGGCTTAAGGACCA     | 449 | GAGACATCTGCGCCCAAGTTAAGCAGCAGCATCAGAAAGGACAGAGTAAACGAGGCCAGG  | 1349 |  |  |  |  |  |
| DB       | 628  | GCTGCTCTCGCTGTGAAGATGATCTTGGAGATTATGATATAAGAGGCTTAAGGACCA     | 687 | GAGACATCTGCGCCCAAGTTAAGCAGCAGCATCAGAAAGGACAGAGTAAACGAGGCCAGG  | 1587 |  |  |  |  |  |
| QY       | 450  | AAACGGCTGACAGGTTCCCTCGAGATCTTGGCCCAATAAATCTTGTAGGGGCTCTCA     | 509 | CACGAGGAAGTGAACAGCCCTCTGTCCTTNGAGTCCAGGCAACCAACGCGGAGAA       | 1409 |  |  |  |  |  |
| DB       | 688  | AAACGGCTGACAGGTTCCCTCGAGATCTTGGCCCAATAAATCTTGTAGGGGCTCTCA     | 747 | CACGAGGAAGTGAACAGCCCTCTGTCCTTNGAGTCCAGGCAACCAACGCGGAGAA       | 1647 |  |  |  |  |  |
| QY       | 510  | TTGGCAAGGAGGAGCGAACCTGAGAGAGGTAGAGCAAGATACCGAGACAAAATCACCA    | 569 | ATCGAGAGTGTGCTCTCCCGCGAGGCTGAGAAATGAGTGGGAATCCCGGACACNTGGGC   | 1469 |  |  |  |  |  |
| DB       | 748  | TTGGCAAGGAGGAGCGAACCTGAGAGAGGTAGAGCAAGATACCGAGACAAAATCACCA    | 807 | ATCGAGAGTGTGCTCTCCCGCGAGGCTGAGAAATGAGTGGGAATCCCGGACACNTGGGC   | 1707 |  |  |  |  |  |
| QY       | 570  | TCTCTCTGTCAGAGCTTACCTTTACACCTTGAGAGACCATCAGTGTGAAGGGG         | 629 | CGGGCTGTAGATCAGGTTTGGCCACTTGATTGAGAAAGATGTTCCAGTGAAGAACCCCTGA | 1529 |  |  |  |  |  |
| DB       | 808  | TCTCTCTGTCAGAGCTTACCTTTACACCTTGAGAGACCATCAGTGTGAAGGGG         | 867 | CGGGCTGTAGATCAGGTTTGGCCACTTGATTGAGAAAGATGTTCCAGTGAAGAACCCCTGA | 1767 |  |  |  |  |  |
| QY       | 630  | CCATCGAGAAATGTTGCGAGGCGGAGCAGGAAATAATGAAGAAAGTTCCGGAGGCGCTATG | 689 | TCINTCAGCCCCAAAACACCCACCCCAATTCGCCCCACACATGTTNGCCCTCGGGGTGTCA | 1589 |  |  |  |  |  |
|          |  |   |     | TCINTCAGCCCCAAAACACCCACCCCAATTCGCCCCACACATGTTNGCCCTCGGGGTGTCA | 1827 |  |  |  |  |  |
|          |  |   |     | AAATTATTAGCGAAGGCACTTTTAAACGTTGATTGTTTAAAGAGCTCTCCAGGCCCCAC   | 1649 |  |  |  |  |  |
|          |  |   |     | AAATTATTAGCGAAGGCACTTTTAAACGTTGATTGTTTAAAGAGCTCTCCAGGCCCCAC   | 1887 |  |  |  |  |  |
|          |  |   |     | CAAGAGGGTGGATCACACCTCAGTGGGAGAGAAATAAATTTCTTCAAGGTTTAAAA      | 1708 |  |  |  |  |  |
|          |  |   |     | CAAGAGGGTGGATCACACCTCAGTGGGAGAGAAATAAATTTCTTCAAGGTTTAAAA      | 1946 |  |  |  |  |  |

AF117106 2130 bp mRNA linear PRI 26-JAN-1999  
LOCUS Homo sapiens IGF-II mRNA-binding protein 1 (IMP-1) mRNA, complete  
DEFINITION cds.  
ACCESSION AF117106  
VERSION AF117106.1 GI:4191607  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 2130)  
AUTHORS Nielsen, J., Christensen, J., Lykke-Andersen, J., Johnsen, A.H.,  
Wewer, U.M. and Nielsen, F.C.  
TITLE A family of insulin-like growth factor II mRNA-binding proteins  
represses translation in late development  
JOURNAL Mol. Cell. Biol. 19 (2), 1262-1270 (1999)  
MEDLINE 99108099  
PubMed 9891060  
REFERENCE 2 (bases 1 to 2130)  
AUTHORS Nielsen, J., Christensen, J., Lykke-Andersen, J., Johnsen, A.H.,  
Wewer, U.M. and Nielsen, F.C.  
TITLE Direct Submission  
JOURNAL Submitted (30-DEC-1998) Institute of Molecular Biology, University  
of Copenhagen, Sølvgade 83H, Copenhagen DK-1307, Denmark  
FEATURES  
source  
location/Qualifiers  
1..2130  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
1..2130  
/gene="IMP-1"  
10..1743  
/gene="IMP-1"  
/note="binds multiple sites in IGF-II mRNA 5' UTR"  
/codon\_start=1  
/product="IGF-II mRNA-binding protein 1"  
/protein\_id="AA009826.1"  
/db\_xref="GI:4191608"  
translation="MNKLYIGNLINESVTPADLEKVFABHKISVGGFLVKSGYAFVDC  
PDEHAKAJETSPGKVELQKLEIEHSVPKQSRKIQIRNPQLREWLDSLLA  
QYGVNCEQVNTSESTAVVNTYSNREQTRQAIMKLNHOLEHAKVSYTPDEQIA  
QPNRGGRGGRGPRGSPVAGAPAKQOVD:PLRLVPTQVVGALIGKEGAT  
RNITQTSKIDVHRKENAGAESAISVHSTPSCSSACKVLEIWHKAKDKTAD  
VPLKLAHNFVGLRIGEGNLIKVEQDPTETKITISSQDLILYNPRTITVKGAL  
NCCRAEIMKKVREAVENDVAAVMSQSLIPGLNLAAGLFFPASAVPPPPSSVTG  
AAVSSPMQPEQEMVQVFIQAQVAGALIGKGQHIKQLSRFASISIKIAPPETPSK  
VRWVITTPPEAQFKAGRIYGLKEENFPKPEVKLETHIRVPASAAGRVIGKSK  
TVNELQNLTAEEVVPRDTPDENDQIVIKIIGHFVASQVQSRKIRDILAQVQHQK  
GQSNQAQARRK"  
ORIGIN  
Query Match 96.8%; Score 1653.2; DB 9; Length 2130;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 1671; Conservative 0; Mismatches 8; Indels 6; Gaps 1;  
Qy 30 CCGGGGAGCCATCATGAAGTGAATGGCCACAGTGGAGAACCATGCCCTGAAGTCT 89  
Db 404 CCAGGCAAGCCATCATGAAGTGAATGGCCACAGTGGAGAACCATGCCCTGAAGTCT 463  
Qy 90 CCTACATCCCGATGACAGATGACAGAGACCTGAGAAATGGCGCCGAGGGGCTTTG 149  
Db 464 CCTACATCCCGATGACAGATGACAGAGACCTGAGAAATGGCGCCGAGGGGCTTTG 523  
Qy 150 GCTCTCGGGTACGCCCGCCAGGGCTCCTCTGGCAGCGGGGGCCCGAGCCACGAC 209  
Db 524 GCCTCGGGTACGCCCGCCAGGGCTCCTCTGGCAGCGGGGGCCCGAGCCACGAC 583  
Qy 210 AGCAAGTGGACATCCCTCTGGTCTCTGGTCCGACCCAGTATGGGTGCCATTATTG 269  
Db 584 AGCAAGTGGACATCCCTCTGGTCTCTGGTCCGACCCAGTATGGGTGCCATTATTG 643  
Qy 270 GCAAGGAGGGGGCCACCACCGCAACATCACAAAAGAGCCAGTCCAGATAGAGCTGC 329

644 GCAGGAGGGGGCCACCATCCGACATCACAAAACAGACCCAGTCCAGATAGACGTGC 703  
Qy 330 ATAGGAAGGAGAACCGCAGGTGCAGTGAAGAACCATCAGTGTGCACCTCCACCTGAGG 389  
Db 704 ATAGGAAGGAGAACCGCAGGTGCAGTGAAGAACCATCAGTGTGCACCTCCACCTGAGG 763  
Qy 390 GCTGCTCTCTCCCTTGTAAAGATGATCTTGGAGATTTATGCATAAAGAGGCTTAAGACACCA 449  
Db 764 GCTGCTCTCTCCCTTGTAAAGATGATCTTGGAGATTTATGCATAAAGAGGCTTAAGACACCA 823  
Qy 450 AAAACGGCTGACAGGTTCCCTGAAGATCTCTGGCCATAAATACTTTAGGGGCTCTCA 509  
Db 824 AAAACGGCTGACAGGTTCCCTGAAGATCTCTGGCCATAAATACTTTAGGGGCTCTCA 883  
Qy 510 TTGGCAAGGAGGAGCGGAACTCTGAAGAGTAGAGATACCGAGACCAAAATACCA 569  
Db 884 TTGGCAAGGAGGAGCGGAACTCTGAAGAGTAGAGATACCGAGACCAAAATACCA 943  
Qy 570 TCTCTCTGTTGCAAGACCTTTACCTTTTACAACCTTGAGAGCACCATCACTGTGAAGGGG 629  
Db 944 TCTCTCTGTTGCAAGACCTTTACCTTTTACAACCTTGAGAGCACCATCACTGTGAAGGGG 1003  
Qy 630 CCATCAGAAATTTGTCAGGCGCCGAGCAGGAAATATGAAGAAAGTTCCGGAGGCTTATG 689  
Db 1004 CCATCAGAAATTTGTCAGGCGCCGAGCAGGAAATATGAAGAAAGTTCCGGAGGCTTATG 1063  
Qy 690 AGAATGATGTGGCTGCCATGAGC-----TCTCACTCTGATCCCTGGCTGAACTGGGTG 743  
Db 1064 AGAATGATGTGGCTGCCATGAGCCTGCTCAGTCTCCTGATCCCTGGCTGAACTGGGTG 1123  
Qy 744 CTGTAGGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCCGCTCCAGCAGCGCTTACTG 803  
Db 1124 CTGTAGGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCCGCTCCAGCAGCGCTTACTG 1183  
Qy 804 GGGCTGCTCCCTATAGTCTCTTTATGAGGCTCCCGAGCAGGATGCTGCGAGTGTGTTA 863  
Db 1184 GGGCTGCTCCCTATAGTCTCTTTATGAGGCTCCCGAGCAGGATGCTGCGAGTGTGTTA 1243  
Qy 864 TCCCGCCAGCAGTGGCGCCCATCATCGCAAGAGGGGCGACACATCAAAACAGTCT 923  
Db 1244 TCCCGCCAGCAGTGGCGCCCATCATCGCAAGAGGGGCGACACATCAAAACAGTCT 1303  
Qy 924 CCGGTTTTGCCAGCGCTCCATCAAGATGCAACACCGCAACCTGACTCCAAAGTTC 983  
Db 1304 CCGGTTTTGCCAGCGCTCCATCAAGATGCAACACCGCAACCTGACTCCAAAGTTC 1363  
Qy 984 GTATGTTTATCATCACTGGACCGCCAGAGGCCCAATTCAAGGCTCAGGAGAAATCTATG 1043  
Db 1364 GTATGTTTATCATCACTGGACCGCCAGAGGCCCAATTCAAGGCTCAGGAGAAATCTATG 1423  
Qy 1044 GCAAACTCAAGAGGAGAACTTTCTTGGTCCCAAGGAGAAAGTGAAGTGGAGACCCACA 1103  
Db 1424 GCAAACTCAAGAGGAGAACTTTCTTGGTCCCAAGGAGAAAGTGAAGTGGAGACCCACA 1483  
Qy 1104 TACGTGTCCAGCATCAGCAGCTGGCGGGTCAATGGCAAGGTGGAAAAACCGTGAACG 1163  
Db 1484 TACGTGTCCAGCATCAGCAGCTGGCGGGTCAATGGCAAGGTGGAAAAACCGTGAACG 1543  
Qy 1164 AGTTGCAAAATTTGACGGCAGCTGAGGTGGTAGTACCAAGAGACCAAGCCCTGATGAGA 1223  
Db 1544 AGTTGCAAAATTTGACGGCAGCTGAGGTGGTAGTACCAAGAGACCAAGCCCTGATGAGA 1603  
Qy 1224 ACGACCGAGTCACTCGTGAATAATCATCGACATTTCTATGCCAGTCAAGTGCCTCAACGGA 1283  
Db 1604 ACGACCGAGTCACTCGTGAATAATCATCGACATTTCTATGCCAGTCAAGTGCCTCAACGGA 1663  
Qy 1284 AGATCCGAGACATCTGGCCCAAGTTTAAAGCAGCAGCATCAAGAGGACAGAGTAACGAG 1343  
Db 1664 AGATCCGAGACATCTGGCCCAAGTTTAAAGCAGCAGCATCAAGAGGACAGAGTAACGAG 1723  
Qy 1344 CCCAGGACCGAGAGTGAACGCCCTCCCTGTCCTTNGAGTCCAGGACACACAGG 1403

Db 1724 CCCAGGCACGAGGAGTGACCGCCCTCCCTGTCCTCCCTGAGTCCAGGACACACACG 1783

Qy 1404 GCAGAAATCGAGATGTGCTCTCCCGGAGGCTGAGATGAGTGGATCCGGACAC 1463

Db 1784 GCAGAAATCGAGATGTGCTCTCCCGGAGGCTGAGATGAGTGGATCCGGACAC 1843

Qy 1464 NTGGCGCGGCTGAGATGAGTGTGCTCTCCCTGAGTGGATGAGTGGATGAGTGGAT 1523

Db 1844 CTGGCGCGGCTGAGATGAGTGTGCTCTCCCTGAGTGGATGAGTGGATGAGTGGAT 1903

Qy 1524 CCCTGATCTNTACGCCCAACACCCACCCCAATTTGGCCCAACACTGNTGCGCTCGGG 1583

Db 1904 CCCTGATCTCTCAGCCCAACACCCACCCCAATTTGGCCCAACACTGNTGCGCTCGGG 1963

Qy 1584 TGTGAGAAATNTAGCCGAGGCACTTTAAACGCTGATTTTAAAGAGCTCTCCAG 1643

Db 1964 TGTGAGAAATNTAGCCGAGGCACTTTTAAACGCTGATTTTAAAGAGCTCTCCAG 2023

Qy 1644 CCCACCAAGAGGCTGATCACCTCAGTGGGAGGAAATAAAATTTCTTCAGGTTT 1703

Db 2024 CCCACCAAGAGGCTGATCACCTCAGTGGGAGGAAATAAAATTTCTTCAGGTTT 2083

Qy 1704 TAAAA 1708

Db 2084 TAAAA 2088

RESULT 8

AF198254 2381 bp mRNA linear PRI 02-MAR-2000

LOCUS AF198254 Homo sapiens mRNA-binding protein CRDBP mRNA, complete cds.

DEFINITION AF198254

ACCESSION AF198254.1 GI:7141071

VERSION

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 2381)

Ioannidis,P., Trangas,T., Dimitriadis,E., Samiotaki,M., Panoutsakopoulos,G., Kyriazoglou,I., Voutzoulas,S., Tsipalis,C.M., Kittas,C., Agnantis,N. and Pandis,N.

Ecopic expression of a KH-domain containing protein, highly homologous to both human IMP-1 and mouse CRD-BP, in benign and malignant mesenchymal tumors

Unpublished

Ioannidis,P., Trangas,T., Dimitriadis,E. and Samiotaki,M. 2 (bases 1 to 2381)

Direct Submission

Submitted (25-OCT-1999) G. Papanikolaou Research Center, St Savvas Hospital, 171 Alexandras Avenue, Athens, Attica 11522, Greece

Location/Qualifiers

1. .2381

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/map="17q21.3-q22"

/cell\_line="MCF-7"

306. .2039

/note="KH-domain containing protein; similar to Mus musculus CRD-BP (c-myc mRNA coding region determinant-binding protein) and Homo sapiens IMP-1 (insulin growth factor II mRNA-binding protein)"

/codon\_start=1

/product="mRNA-binding protein CRDBP"

/protein\_id="AF37203.1"

/db\_xref="GI:7141072"

/translation="MNKLYIGNINESVTPADLEKVFAEHKISYSGQFLVKGAVFVDC PDERWAMKALETFSKVELQKRLIEHSVPKQSRKIQIRNPOLRWEVLDSLIA QGTGENCEQNTSEAVNTYSNREQTQALMKLGHQLENHALKVSYPDEQIA QGPENRGSGQSGPQVPAAGAKQOQVDIPRLRLVPTQYVGLIGEGATI RNITQTQSKIDVHRKENAGAAEKAISVHSTPEGSSACKRMILEIMHEAKDITKTADE

FEATURES

source

CDS

ORIGIN

Query Match 95.5%; Score 1631.8; DB 9; Length 2381; Best Local Similarity 98.9%; Pred. No. 0; Matches 1663; Conservative 0; Mismatches 12; Indels 7; Gaps 2;

Qy 30 CCCGGGAGCCATCATGAACTGAAATGGCCACCACTTGGAGAACCATGCCCTGAAGTCT 89

Db 700 CCAGGCAAGCCATCATGAACTGAAATGGCCACCACTTGGAGAACCATGCCCTGAAGTCT 759

Qy 90 CCTACATCCCCCATGAGCATAGCACAGGAGCTGAGAACTGGCGCGCGAGGGGCTTTG 149

Db 760 CCTACATCCCCCATGAGCATAGCACAGGAGCTGAGAACTGGCGCGCGAGGGGCTTTG 819

Qy 150 GCTCTCGGGGTGAGCGCCCGCCAGGGCTCACCTGTGGCAGCGGGGGCCCCAGCAAGCAGC 209

Db 820 GCTCTCGGGGTGAGCGCCCGCCAGGGCTCACCTGTGGCAGCGGGGGCCCCAGCAAGCAGC 879

Qy 210 AGCAAGTGGACATCCCCCTTCGGTCTCGGTGCGCCACCCAGTATGTGGTGCATTTATG 269

Db 880 AGCAAGTGGACATCCCCCTTCGGTCTCGGTGCGCCACCCAGTATGTGGTGCATTTATG 939

Qy 270 GAAAGAGGGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCCAAAGATAGACGTGC 329

Db 940 GCNAGAGGGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCCAAAGATAGACGTGC 999

Qy 330 ATAGGAAGGAGAACGCGAGTGCAGCTGAAAGAGCCATCATGTGTGCATCTCACCCCTGAGG 389

Db 1000 ATAGGAAGGAGAACGCGAGTGCAGCTGAAAGAGCCATCATGTGTGCATCTCACCCCTGAGG 1059

Qy 390 GCTGCTCCTCCGCTTCTGAGATGATCTTGGAGATATGATCAAAAGAGGTAAAGGACACCA 449

Db 1060 GCTGCTCCTCCGCTTCTGAGATGATCTTGGAGATATGATCAAAAGAGGTAAAGGACACCA 1119

Qy 450 AAAAGGCTGACGAGGTTCCTCCCTGAGATCCTCGGCCCAATAATACTTTGTAGGCGCTCA 509

Db 1120 AAAAGGCTGACGAGGTTCCTCCCTGAGATCCTCGGCCCAATAATACTTTGTAGGCGCTCA 1179

Qy 510 TTGGCAAGGAGAACGCGAGTGCAGAACGAGTGCAGAACGAGTGCAGAACGAGTGCAGAAC 569

Db 1180 TTGGCAAGGAGAACGCGAGTGCAGAACGAGTGCAGAACGAGTGCAGAACGAGTGCAGAAC 1239

Qy 570 TCTCCTCGTTGCAAGACCTTACCTTTTACAACTGAGAGGACCATCATCTGTGAAGGGGG 629

Db 1240 TCTCCTCGTTGCAAGACCTTACCTTTTACAACTGAGAGGACCATCATCTGTGAAGGGGG 1299

Qy 630 CCATCGAGAAATTTGTCAGGGGCGAGCAGAGAAATAATGAAGAAAGTTTCGGAGGCGCTATG 689

Db 1300 CCATCGAGAAATTTGTCAGGGGCGAGCAGAGAAATAATGAAGAAAGTTTCGGAGGCGCTATG 1359

Qy 690 AGAATGATGTGGTGCATGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 743

Db 1360 AGAATGATGTGGTGCATGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1419

Qy 744 CTGTAGTCTTTTCCAGGCTTTCATCCAGGCGAGTCCCGCGCGCTCCCGAGCGGTTACTG 803

Db 1420 CTGTAGTCTTTTCCAGGCTTTCATCCAGGCGAGTCCCGCGCGCTCCCGAGCGGTTACTG 1479

Qy 804 GGGCTGTCTCCCTATAGCTCTCTTTATCAGGCTCCCGAGCAGGAGTGTGCGAGTCTTTA 863

Db 1480 GGGCTGTCTCCCTATAGCTCTCTTTATCAGGCTCCCGAGCAGGAGTGTGCGAGTCTTTA 1539

Qy 864 TCCCGGCGCAGGAGTGGGGCGCCATCATCGGCAAGAGGGGCGAGCAGCAGCAGCAGCAGCTCT 923

Db 1540 TCCCGGCGCAGGAGTGGGGCGCCATCATCGGCAAGAGGGGCGAGCAGCAGCAGCAGCTCT 1599

Qy 924 CCGGGTTTTCAGGCGCTCCATCAAGATTGCAACCCCGAAACACCTGACTCTCCAAAGTTC 983

```
Db 1600 CCCGTTTGCAGCGCTCCATCAAGATTGCACCCGGAACACCTGACTCCAAAGTTC 1659
Qy 984 GTATGTTATCATCTAGCGCCGAGAGGCCCAATTCAGGCTCAGGGAAGAATCTATG 1043
Db 1660 GTATGTTATCATCTAGCGCCGAGAGGCCCAATTCAGGCTCAGGGAAGAATCTATG 1719
Qy 1044 GCARACTCAAGAGAGACTCTTTGGTCCAGAGGAAGTGAAGCTGAGACCCACA 1103
Db 1720 GCARACTCAAGAGAGACTCTTTGGTCCAGAGGAAGTGAAGCTGAGACCCACA 1779
Qy 1104 TACGTGTGCAGCATCAGCAGCTGCGCGGTCATTGGCAAAAGGTGGAAGGCGTGAACG 1163
Db 1780 TACGTGTGCAGCATCAGCAGCTGCGCGGTCATTGGCAAAAGGTGGAAGGCGTGAACG 1839
Qy 1164 AGTTCAGAAATTGACGAGCTGAGTGGTAGTACCAAGAGACAGACCCCTGATGAGA 1223
Db 1840 AGTTCAGAAATTGACGAGCTGAGTGGTAGTACCAAGAGACAGACCCCTGATGAGA 1899
Qy 1224 AGCAGCAGTCACTCGTGAATCATCGGACATTTCTATGCCAGTCAGATGCTCAACGGA 1283
Db 1900 AGCAGCAGTCACTCGTGAATCATCGGACATTTCTATGCCAGTCAGATGCTCAACGGA 1959
Qy 1284 AGATCCGAGACATCTGCGCCAGGTTAAGCAGCAGCATCAGAAGGGAAGAGTAAACAGG 1343
Db 1960 AGATCCGAGACATCTGCGCCAGGTTAAGCAGCAGCATCAGAAGGGAAGAGTAAACAGG 2019
Qy 1344 CCCAGGACGAGGAAGTACCGCCCTCCCTGCTCCCTTNGAGTCCAGGACACACACG 1403
Db 2020 CCCAGGACGAGGAAGTACCGCCCTCCCTGCTCCCTTNGAGTCCAGGACACACACG 2079
Qy 1404 CGAGAAATCGAGAGTGTCTCTCCCGGAGGCTGAGAAATGAGTGGGAATCCGGGACAC 1463
Db 2080 CGAGAAATCGAGAGTGTCTCTCCCGGAGGCTGAGAAATGAGTGGGAATCCGGGACAC 2139
Qy 1464 NTGGCGGGGTGAGATCAGGTTTGCACCTTTCAGTTGAGAAAGATGTTCCAGTGAAGAA 1523
Db 2140 CTGGCGGGGTGAGATCAGGTTTGCACCTTTCAGTTGAGAAAGATGTTCCAGTGAAGAA 2199
Qy 1524 CCCTGATCTNTCAG-CCCCAAACACCCCAATGCGCCCAACACTGNTGCCCTCGGG 1582
Db 2200 CCCTGATCTCTCGCCCGGAGGCTTACCAATTTGCCCAACACTGCTGCCCTCGGG 2259
Qy 1583 GTGTGAGAAATTAGCGCAAGGACCTTTAAAGTGGATGTTTAAAGAGTCTCCAG 1642
Db 2260 GTGTGAGAAATTAGCGCAAGGACCTTTTAAAGTGGATGTTTAAAGAGTCTCCAG 2319
Qy 1643 GCCCCACCAAGAGGTTGATCAGCTTTCAGTGGGGAAGAAATAAATTTCCCTTCAGTT 1702
Db 2320 GCCCCACCAAGAGGTTGATCAGCTTTCAGTGGGGAAGAAATAAATTTCCCTTCAGTT 2379
Qy 1703 TT 1704
Db 2380 TT 2381

RESULT 9
BD127811
LOCUS 2780 bp DNA linear PAT 18-SEP-2002
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD127811
VERSION BD127811.1 GI:23222756
KEYWORDS JP 2002017375-A/3242.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2780)
Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Koga,H.
Primer for synthesizing full-length cDNA and use thereof
JOURNAL Primer: JP 2002017375-A 3242 22-JAN-2002;
```

```
COMMENT HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002017375-A/3242
PD 22-JAN-2002
PF 07-JUL-2000 JP 20020253172
PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO
PI ISHII,
PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI
SHINICHI KOJIMA,
PI TETSUJI OTSUKI, HISASHI KOGA
PC
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10, C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers
(112)..(1068).
FT CDS
FEATURES
Source
1..2780
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 81.3%; Score 1389; DB 6; Length 2780;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1405; Conservative 0; Mismatches 5; Indels 6; Gaps 1;
Qy 299 ACAAACAGACCCAGTCCAAAGATAGACGTGCATAGAGAGGAGAAACGAGGTGCAGTCAA 358
Db 1 ACAAACAGACCCAGTCCAAAGATAGACGTGCATAGAGAGGAGAAACGAGGTGCAGTCAA 60
Qy 359 AAAGCCATCAGTGCACCTCCACCCCTGAGGCTCTCCCTCGCTTGAAGATGATCTTG 418
Db 61 AAAGCCATCAGTGCACCTCCACCCCTGAGGCTCTCCCTCGCTTGAAGATGATCTTG 120
Qy 419 GAGATTATGCATAAAGAGGCTAAGGACACCAAAACGGCTGACGAGGTTCCCTCAAGATC 478
Db 121 GAGATTATGCATAAAGAGGCTAAGGACACCAAAACGGCTGACGAGGTTCCCTCAAGATC 180
Qy 479 CTGGCCCATATTAATCTTTAGGGGCTCTCATTTGCCAAGAGAGGACGAACTGAAGAG 538
Db 181 CTGGCCCATATTAATCTTTAGGGGCTCTCATTTGCCAAGAGAGGACGAACTGAAGAG 240
Qy 539 GTAGAGCAAGATACCGAGACAAAATCAACATCTCTCGTTGCAAGACCTTACCCTTTAC 598
Db 241 GTAGAGCAAGATACCGAGACAAAATCAACATCTCTCGTTGCAAGACCTTACCCTTTAC 300
Qy 599 AACCTGAGAGACCATCAGTGTGAAGGGGCGCATCGAGAAATTTGTCAGGGCGGAGAG 658
Db 301 AACCTGAGAGACCATCAGTGTGAAGGGGCGCATCGAGAAATTTGTCAGGGCGGAGAG 360
Qy 659 GAAATAATGAAGAAAGTTCCGGAGGCGCTATGAGAAATGATGTGGCTGCCATGAGC- 712
Db 361 GAAATAATGAAGAAAGTTCCGGAGGCGCTATGAGAAATGATGTGGCTGCCATGAGGCTGAG 420
Qy 713 TCTCACTGATCCCTGGGCTGAACTGGCTGCTGAGTGTCTTTCCAGTGTTCATCCAGC 772
Db 421 TCTCACTGATCCCTGGGCTGAACTGGCTGCTGAGTGTCTTTCCAGTGTTCATCCAGC 480
Qy 773 GCAGTCCCGCGGCTCCAGCAGAGGTTACTTGGGGCTGCTCCCTAGCTCTCTTTATCAG 832
Db 481 GCAGTCCCGCGGCTCCAGCAGAGGTTACTTGGGGCTGCTCCCTAGCTCTCTTTATCAG 540
Qy 833 GCTCCCGAGCAGGAGATGTTGAGTGTGTTATCCCGCCAGGAGTGGGCGGATCATC 892
Db 541 GCTCCCGAGCAGGAGATGTTGAGTGTGTTATCCCGCCAGGAGTGGGCGGATCATC 600
Qy 893 GGCAGAGAGGGGAGCAGCAGTCCAGCAGAGGTTACTTGGGGCTGCTCCCTAGCTCTCTTTATCAG 952
Db 601 GGCAGAGAGGGGAGCAGCAGTCCAGCAGAGGTTACTTGGGGCTGCTCCCTAGCTCTCTTTATCAG 660
Qy 953 GCACCAACCCGAAACACCTGCTCAAAAGTTGATGTGTTATCATCACTGAGCAGCCAGAG 1012
```

|            |   |   |      |
|------------|---|---|------|
| Db         | 661   | GCACACCCGAAACACCTGACTCCAAAGTTCTGTATGTTATCATCTGGACCCGACAG        | 720  |
| Qy         | 1013  | GCCCAATTCAAGGCTCAGGAGAACTCTATGCAAACTCAAGGAGGAGAACTCTTTGGT       | 1072 |
| Db         | 721   | GCCCAATTCAAGGCTCAGGAGAACTCTATGCAAACTCAAGGAGGAGAACTCTTTGGT       | 780  |
| Qy         | 1073  | CCCAAGGAGGAGTGAAGTGGAGAGCCACATAGTGTGCCAGCATCAGAGCTGGCGG         | 1132 |
| Db         | 781   | CCCAAGGAGGAGTGAAGTGGAGAGCCACATAGTGTGCCAGCATCAGAGCTGGCGG         | 840  |
| Qy         | 1133  | GTCAATTTGCAAAAGGTGGAAAAACGGTGAACGAGTTCAGATTTCAGCGGACGCTGAGTG    | 1192 |
| Db         | 841   | GTCAATTTGCAAAAGGTGGAAAAACGGTGAACGAGTTCAGATTTCAGCGGACGCTGAGTG    | 900  |
| Qy         | 1193  | GTAGTACCAAGAGACCAAGCCCTGATGAGAACGACGACAGGTCTATCGTGAATAATCATCGGA | 1252 |
| Db         | 901   | GTAGTACCAAGAGACCAAGCCCTGATGAGAACGACGACAGGTCTATCGTGAATAATCATCGGA | 960  |
| Qy         | 1253  | CATTTCTATGTCAGTTCAGATGCTCAACGGAGAGTCCGAGACATCTTGGCCAGGTTAAG     | 1312 |
| Db         | 961   | CATTTCTATGTCAGTTCAGATGCTCAACGGAGAGTCCGAGACATCTTGGCCAGGTTAAG     | 1020 |
| Qy         | 1313  | CAGCAGCATCAGAAAGGACAGAGTAACCAAGCCCGAGGACCGAGGAAAGTGACCAAGCCCT   | 1372 |
| Db         | 1021  | CAGCAGCATCAGAAAGGACAGAGTAACCAAGCCCGAGGACCGAGGAAAGTGACCAAGCCCT   | 1080 |
| Qy         | 1373  | CCCTGTCCCTTNGAGTCCAGGACACCAAGCGGACGAATCCAGAGTGTGCTTCCCGGC       | 1432 |
| Db         | 1081  | CCCTGTCCCTTNGAGTCCAGGACACCAAGCGGACGAATCCAGAGTGTGCTTCCCGGC       | 1140 |
| Qy         | 1433  | AGGCTGAGATGAGTGGGAATCCGGGACACNTGGCCGGCTGTAGATCAGGTTTGCC         | 1492 |
| Db         | 1141  | AGGCTGAGATGAGTGGGAATCCGGGACACCTGGCCGGCTGTAGATCAGGTTTGCC         | 1200 |
| Qy         | 1493  | ACTTGATTGAGAAAGATGTTCCAGTGAAGAACCTTGATCTWTAGCCCAACCAACCCACC     | 1552 |
| Db         | 1201  | ACTTGATTGAGAAAGATGTTCCAGTGAAGAACCTTGATCTWTAGCCCAACCAACCCACC     | 1260 |
| Qy         | 1553  | CAATTGGCCCAACACTGNTCCCTCGGGGTGTCAGAAATTTAGCGCAGGACATTTT         | 1612 |
| Db         | 1261  | CAATTGGCCCAACACTGNTCCCTCGGGGTGTCAGAAATTTAGCGCAGGACATTTT         | 1320 |
| Qy         | 1613  | AAAGTGTGATTGTTTAAAGAGCTCTCAGGCCCCCACAAGAGGTTGATCACAACCTCAG      | 1672 |
| Db         | 1321  | AAAGTGTGATTGTTTAAAGAGCTCTCAGGCCCCCACAAGAGGTTGATCACAACCTCAG      | 1380 |
| Qy         | 1673  | TGGGAGGAAATAAATAATTTCTTCAGGTTTAAAA                              | 1708 |
| Db         | 1381  | TGGGAGGAAATAAATAATTTCTTCAGGTTTAAAA                              | 1416 |
| RESULT 10  |   |   |      |
| AK074915   |   |   |      |
| LOCUS      | AK074915 2780 bp mRNA linear PRI 03-SEP-2002  |   |      |
| DEFINITION | Homo sapiens cDNA FLJ90434 fis, clone NT2RP3000789, highly similar to Homo sapiens IGF-II mRNA-binding protein 1 (IMP-1) mRNA.  |   |      |
| ACCESSION  | AK074915  |   |      |
| VERSION    | AK074915.1 GI:22760672  |   |      |
| KEYWORDS   | oligo capping; fis (full insert sequence).  |   |      |
| SOURCE     | Homo sapiens (human)  |   |      |
| ORGANISM   | Homo sapiens  |   |      |
| REFERENCE  | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.   |   |      |
| AUTHORS    | 1 Isogai, T., Ota, T., Nishikawa, T., Hayashi, K., Sugano, S., Ishii, S., Sugiyama, T., Suzuki, Y., Nagai, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Kawai-Hio, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Kojima, S., Nagahara, K., Masuho, Y., Ono, T., Okano, K., Yoshikawa, Y., Aotsuka, S., Sasaki, N., Hattori, A., Okumura, K., Iwayanagi, T. and Ninomiya, K. |   |      |
| TITLE      | NEDO human cDNA sequencing project  |   |      |
| JOURNAL    | Unpublished   |   |      |

|           |  |  |                                 |
|-----------|--|--|---------------------------------|
| REFERENCE | 2 (bases 1 to 2780)  |  |                                 |
| AUTHORS   | Isogai, T. and Otsuki, T.  |  |                                 |
| TITLE     | Direct Submission  |  |                                 |
| JOURNAL   | Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986) NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction: Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; cDNA 5' - & 3' - end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.). |  |                                 |
| FEATURES  | Location/Qualifiers  |  |                                 |
| source    | 1. 2780  |  |                                 |
|           | /organism="Homo sapiens"   |  |                                 |
|           | /mol_type="mRNA"   |  |                                 |
|           | /db_xref="taxon:9606"  |  |                                 |
|           | /clone="NT2RP3000789"  |  |                                 |
|           | /cell_line="NT2"   |  |                                 |
|           | /cell_type="teratocarcinoma"   |  |                                 |
|           | /clone_lib="NT2RP3"  |  |                                 |
|           | /notes="cloning vector: pME18SFL3-mRNA from NT2 neuronal precursor cells after 2-weeks retinoic acid (RA) induction"   |  |                                 |
| ORIGIN    |  |  |                                 |
|           | Query Match  | 81.3%;   | Score 1389; DB 9; Length 2780;  |
|           | Best Local Similarity  | 99.2%;   | Pred. No. 0;                    |
|           | Matches 1405; Conservative   | 0;   | Mismatches 5; Indels 6; Gaps 1; |
| Qy        | 299  | ACAAAACAGACCCAGTCCCAAGATAGACGTGATAGGAAGAGAACCCAGGTGAGCTGAA     | 358                             |
| Db        | 1  | ACAAAACAGACCCAGTCCCAAGATAGACGTGATAGGAAGAGAACCCAGGTGAGCTGAA     | 60                              |
| Qy        | 359  | AAGCCATCAGTGTGCACTCCACCCCTGAGGGTGTCTCTCCGCTTGTAGATGATCTTG      | 418                             |
| Db        | 61   | AAGCCATCAGTGTGCACTCCACCCCTGAGGGTGTCTCTCCGCTTGTAGATGATCTTG      | 120                             |
| Qy        | 419  | GAGATTATCATAAAGAGGCTAAGGACACAAACGGCTGACGAGGTTCCTCCGTAAGATC     | 478                             |
| Db        | 121  | GAGATTATCATAAAGAGGCTAAGGACACAAACGGCTGACGAGGTTCCTCCGTAAGATC     | 180                             |
| Qy        | 479  | CTGGCCCATATAAATCTTTGTAGGGCTCTCATTTGGCAAGGAAGGACGGAACCTGAAAG    | 538                             |
| Db        | 181  | CTGGCCCATATAAATCTTTGTAGGGCTCTCATTTGGCAAGGAAGGACGGAACCTGAAAG    | 240                             |
| Qy        | 539  | GTAGACAAGATACCCAGACAAAATCAGCATCTCTCGTTGCAAGACCTTACCTTTTAC      | 598                             |
| Db        | 241  | GTAGACAAGATACCCAGACAAAATCAGCATCTCTCGTTGCAAGACCTTACCTTTTAC      | 300                             |
| Qy        | 599  | AACCTCTGAGAGGACCATCACTGTGAAGGGGGCCATCGAGAATTGTTGCAAGGGCCGAGCAG | 658                             |
| Db        | 301  | AACCTCTGAGAGGACCATCACTGTGAAGGGGGCCATCGAGAATTGTTGCAAGGGCCGAGCAG | 360                             |
| Qy        | 659  | GAAATATGAGAAAGTTCCGGAGGCTATGAGAATGATGTGGCTGCCATGAGC-----       | 712                             |
| Db        | 361  | GAAATATGAGAAAGTTCCGGAGGCTATGAGAATGATGTGGCTGCCATGAGCCTGCGAG     | 420                             |
| Qy        | 713  | TCTCAGCTGATCCCTGGCTGAACTGCTGTAGGTCTTTTCCAGCTTATCATCCAGC        | 772                             |
| Db        | 421  | TCTCAGCTGATCCCTGGCTGAACTGCTGTAGGTCTTTTCCAGCTTATCATCCAGC        | 480                             |
| Qy        | 773  | GCAAGTCCCGCGCTCCAGCAGCGTTACTGGGGCTGTCTCCCTATAGCTCTCTTTATGCG    | 832                             |
| Db        | 481  | GCAAGTCCCGCGCTCCAGCAGCGTTACTGGGGCTGTCTCCCTATAGCTCTCTTTATGCG    | 540                             |
| Qy        | 833  | GCTCCCGAGCAGGAGTGTGAGGTGTTTATCCCGCCAGCAGTGGCGGCCATCATC         | 892                             |
| Db        | 541  | GCTCCCGAGCAGGAGTGTGAGGTGTTTATCCCGCCAGCAGTGGCGGCCATCATC         | 600                             |
| Qy        | 893  | GGCAAGAGGGGAGCAGCATCAACAGCTCTCCCGGTTTGGCAGCGCTCCATCAAGATT      | 952                             |

Db 601 GGCAAGAGGGGAGCAGCATCAACAGAGCTCTCCGGTTTGCCAGGGCTCCATCAAGATT 660  
Qy 953 GCACACACCGGAACACCTGACTCCAAAGTTCGTATGGTTATCATCACTGAGCCGCCAGAG 1012  
Db 661 GCACACACCGGAACACCTGACTCCAAAGTTCGTATGGTTATCATCACTGAGCCGCCAGAG 720  
Qy 1013 GCCCAATTCAAGGCTCAGGGAAGATCTATGGCAAACTCAAGAGGAGAACTCTTTTGGT 1072  
Db 721 GCCCAATTCAAGGCTCAGGGAAGATCTATGGCAAACTCAAGAGGAGAACTCTTTTGGT 780  
Qy 1073 CCCAAGAGGAGTGAAGCTGGAGCCACATATAGTGTGCCAGCATCAGAGCTTGGCCGG 1132  
Db 781 CCCAAGAGGAGTGAAGCTGGAGCCACATATAGTGTGCCAGCATCAGAGCTTGGCCGG 840  
Qy 1133 GTCAATTGGCAAGGTGGAAAAACGGTGAACGAGTTTGAGAAATTTGACGGCAGCTGAGGTG 1192  
Db 841 GTCAATTGGCAAGGTGGAAAAACGGTGAACGAGTTTGAGAAATTTGACGGCAGCTGAGGTG 900  
Qy 1193 GTAGTACCAAGAGACAGACCCCTGATGAGAACGACAGAGTCATCTGTGAAATCATCGGA 1252  
Db 901 GTAGTACCAAGAGACAGACCCCTGATGAGAACGACAGAGTCATCTGTGAAATCATCGGA 960  
Qy 1253 CATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCCGAGACATCTCGGCCAGGTTAAG 1312  
Db 961 CATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCCGAGACATCTCGGCCAGGTTAAG 1020  
Qy 1313 CAGCAGCATCAAGAGGAGCAGAGTAAACAGCCCGAGGACGAGGAAAGTGAACAGCCCT 1372  
Db 1021 CAGCAGCATCAAGAGGAGCAGAGTAAACAGCCCGAGGACGAGGAAAGTGAACAGCCCT 1080  
Qy 1373 CCTCTCCTCTTNGAGTCAGAGACACAGCCGAGAGAAATCGAGAGTGTCTCTCCCGGC 1432  
Db 1081 CCTCTCCTCTTNGAGTCAGAGACACAGCCGAGAGAAATCGAGAGTGTCTCTCCCGGC 1140  
Qy 1433 AGGCTTGAGAAATCGGGGACACNTGGCGGGCTGTAGATCAGGTTTGGCC 1492  
Db 1141 AGGCTTGAGAAATCGGGGACACNTGGCGGGCTGTAGATCAGGTTTGGCC 1200  
Qy 1493 ACTTGATGAGAAAGATTTTCAGTGAGGAAACCTGATCTNTAGCCGCCAAACACCCACC 1552  
Db 1201 ACTTGATGAGAAAGATTTTCAGTGAGGAAACCTGATCTNTAGCCGCCAAACACCCACC 1260  
Qy 1553 CAATTGGCCCAACACTGNTCCCTCGGGGTGTCAGAAATNTAGCCGCAAGGCACTTTT 1612  
Db 1261 CAATTGGCCCAACACTGNTCCCTCGGGGTGTCAGAAATNTAGCCGCAAGGCACTTTT 1320  
Qy 1613 AAACGTGATGTTTAAAGAGCTCTCAGGCCCCACCAAGAGGAGTGGATCACAACCTCAG 1672  
Db 1321 AAACGTGATGTTTAAAGAGCTCTCAGGCCCCACCAAGAGGAGTGGATCACAACCTCAG 1380  
Qy 1673 TGGGAAGAAAAATAAATTTCCCTTCAGGTTTAAAA 1708  
Db 1381 TGGGAAGAAAAATAAATTTCCCTTCAGGTTTAAAA 1416

RESULT 11  
LOCUS AR160244 2224 bp DNA linear PAT 17-OCT-2001  
DEFINITION Sequence 1 from patent US 6255055.  
ACCESSION AR160244  
VERSION AR160244.1 GI:16223887  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE  
1 (bases 1 to 2224)  
AUTHORS Ross,J.  
TITLE c-myc coding region determinant-binding protein (CRD-BP) and its  
nucleic acid sequence  
JOURNAL Patent: US 6255055-A 1 03-JUL-2001;  
FEATURES Location/Qualifiers  
source 1..2224  
/organism="unknown"

ORIGIN  
Query Match 68.2%; Score 1165; DB 6; Length 2224;  
Best Local Similarity 82.8%; Pred. No. 1.8e-282;  
Matches 1411; Conservative 0; Mismatches 260; Indels 33; Gaps 6;  
/mol\_type="unassigned DNA"  
Qy 30 CCGGGGAGCCATCATGAAGCTGAATGGCCACCACTTGGAGAACCATCCCTGGAAGGTCT 89  
Db 525 CCAGGCAAGCTATCATGAAGCTAATGGCCATCACTTGAGAACCATCCCTGGAAGGTCT 584  
Qy 90 CTATCATCCCCATGAGCAGATAGCAGAGGACCTTGAGAAATGGGGCCGAGGGGCTTTG 149  
Db 585 CTTACATATCTGATGAGCAGATAACACAAAGGTCTCTGAGAATGGGGCTCTGAGAGCTTTG 644  
Qy 150 GTCTCGGGGTGAGCCCGCCAGGGCTCACCTGTGGCAGCGGGGGCCCCAGCAAGCAGC 209  
Db 645 GGTCTCGGGGCCAGCCCGGCAAGGTGCGCCGTGGCAGCAGGGCTCCAGCCAGCAGC 704  
Qy 210 AGCAAGTGGACATCCCTCTTGGCTTCCTGGTCCCAACCCAGTATGTGGGTGCATATTG 269  
Db 705 AGCCAGTGGACATCCCTCTTGGCTTCCTGGTTCCTGAGTATGTAGCGGTATCATTTG 764  
Qy 270 GCAAGGAGGGGCGACCATCCGACATCACAAACAGACCCAGTCCAAAGATAGACGTGC 329  
Db 765 GCAAGGAGGGTGCACCATCCGAAACATCACAAACAGACGCGATCCAAATAGACGTGC 824  
Qy 330 ATAGGAAGAGAGAACGAGTGCAGCTGAAAAGCCATCAGTGTGCATCCACCCCTGAGG 389  
Db 825 ATAGGAAGAGAGAACGAGTGCAGCTGAAAAGCCATCAGCGTGCATTCACCCCTGAAG 884  
Qy 390 GTGTCTCTCCCTTGTAGATGATCTTGGAGATTTATGCAATTAAGAGGCTAAGACACA 449  
Db 885 GTGTCTCTCCCTGCGAGATGATCTTGGAGATTTATGCAAGAGGAGCAAGACACA 944  
Qy 450 AAACCGCTGACAGAGTTCCTCTGAAAGATCTCTGGCCCAATAATACTTTAGGGGTCTCA 509  
Db 945 AAACCGCAGTGAAGTTCCTCTGAAAGATCTCTGGCTCTAAACACTTCCTCGGGCGACTCA 1004  
Qy 510 TTGGCAAGAGAGGACGGAACCTGAAAGAGGTAGACAAAGATACCGAGACAAATACACA 569  
Db 1005 TTGGCAAGAGAGGCGGGAACCTGAAAGAGGTGGAGCAGACACAGAGACGAAGATCACA 1064  
Qy 570 TCTCTCGTGTGCAAGACCTTACCTTTTAAACCTGAGAGACCATCACTGTGAAGGGGG 629  
Db 1065 TCTCTCGTGTGCAAGACCTTACCTTTTAAACCTGAGAGACCATCACTGTGAAGGGGG 1124  
Qy 630 CCATCGAAGTGTGCAAGGCGGAGCAGGAAATTAATAAGAAAGTTCCGGAGGCGCTATG 689  
Db 1125 CCATTGAGAACTGTGTCAGGGCGGAGCAGAGATCATGAAGAAAGTTCCAGAGGCTTACG 1184  
Qy 690 AGAATGATGTGCTGCCATGAGC-----TCTCACTCATCCCTGGCTGAACCTGGGTG 743  
Db 1185 AGAAGAGGTGCGCCCATGAGCTTGAGTCCCATCTCATCCCTGGGTGTTAACTGGGTG 1244  
Qy 744 CTGTAGGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCCCGCTCCAGAGCGCTTACTG 803  
Db 1245 CTGTAGGTCTTTTCCAGCTTCATCCAGCGCTGCTCCCTCTCTCCAGCAGTGTCACTG 1304  
Qy 804 GGGCTGCTCCCTATAGCTCTTTATGAGGCTCCCGAGCAGGAGATGTGTGAGGTGTTTA 863  
Db 1305 GGGCTGCTCCCTATAGCTCTTTATGAGGCTCCCGAGCAGGAGATGTGTGAGGTGTTTA 1364  
Qy 864 TCCCGCGCCAGGAGTGGGGCGCCATCATCGGCAAGAGGGGAGCAGACATCAAAAGCTCT 923  
Db 1365 TCCCGCGCCAGGCTGTGGGGCCATCATTTGGCAAGAGGGCCAGCAGATCAACACTCT 1424  
Qy 924 CCGGTTTGCAGCGCTTCCATCAAGATGACACCCGAAACCTGACTGACTCAAGTTC 983  
Db 1425 CCGGTTTGCAGCGCTTCCATCAAGATGACTGACTGACTGACTGACTGACTGACTGACT 1484  
Qy 984 GTATGTTATCATCACTGAGCGCCAGAGGGCCCAATTCAGGGCTCAGGGAAGATCTATG 1043



```
Db 1485 GAATGTCGTCTACTGAGACCCCGAGAGGCTCAGTTCAAGGCTCAGGGAAGATTATG 1544
QY 1044 GCAAACTCAAGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTGAAGTGGAGACCCACA 1103
Db 1545 GCAAACTCAAGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTGAAGTGGAGACCCACA 1604
QY 1104 TACGTGTCCAGCATCAGAGTGGCGGCTTCATTTGGTCCCAAGGAGGAAGTGAAGTGGAG 1163
Db 1605 TACGGGTTCCGGCTTCAGAGCGCGCGCTCATCGGCAAGGCGGCAAGTGAAGTGAAGT 1664
QY 1164 AGTTGAGAAATTTGACGGCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1223
Db 1665 AGCTGAGAACTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1724
QY 1224 AGACAGAGTATCTGAAATCATCGGACATTTCTTATGCGAGTGAAGTGAAGTGAAGTGA 1283
Db 1725 AGACAGAGTATCTGAAATCATCGGACATTTCTTATGCGAGTGAAGTGAAGTGAAGTGA 1784
QY 1284 AGATCCGAGACATCTGCGCCAGGTTAAGCAGCAGCATCAGAGGAGGAGGAGGAGGAGG 1343
Db 1785 AGATCCGAGACATCTGCGCTAAGTTAAGCAGCAGCATCAGAGGAGGAGGAGGAGGAGG 1844
QY 1344 CCCAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1402
Db 1845 CCCAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1904
QY 1403 GGCAGAA-----ATCGAGAGTGTCTCTCCCGGCGAGGCTGAG 1441
Db 1905 AGGAACACAGAACTGGAGGGCGGCTGAGAGGCGCGCTCTCTCCCGGCGAGGCTGAG 1964
QY 1442 AATGAGTGGGAATCCGGACACNTGGCGCGGCTGTAGATCAGGTTTCCCGCACTGATTG 1501
Db 1965 AATGAGTGGGAATCAGG-CATTTGGCGTGGCTGGAGATCAGGTTTGCACACTGATTG 2023
QY 1502 AGAAGAGTTCACAGTGAAGACCTGATCTNTAGCCGCCAAACACCCACCAATGCGCC 1561
Db 2024 AGAACAATGTTCCAGTGAAGATCTGATCTCTCGCCGCCAA--TTGAGCCAGCTGGCCA 2081
QY 1562 CAACACTGNTGCCCTCGGGGTGCAGAAATTTAGCGCAAGGCACTTTTAAACGTGGA 1621
Db 2082 CAGCCCCACCTTGGAAATATCACCATTGCAATCATAGTTGGTGTCTTTTAAACGTGGA 2141
QY 1622 TTGTTTAAAGAGTCTTCAGGCCCCCAAGAGGAGGTTGATCAGCTCAGTGGGAAGAA 1681
Db 2142 TTGCTCT--TGAAGTCTCCAGCTCCATGGAAGGATGGTGCAGATCCAGTCCAGTGGGAAG 2199
QY 1682 AAATAAATTTCTTCAGGTTTAA 1705
Db 2200 AAATAAATTTCTTCAGGTTTAA 2223

RESULT 12
AF061569
LOCUS Mus musculus coding region determinant binding protein mRNA,
DEFINITION complete cds.
ACCESSION AF061569
VERSION AF061569.1 GI:3273748
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2223)
AUTHORS Bernstein,P.L., Herrick,D.J., Prokipcak,R.D. and Ross,J.
TITLE Control of c-myc mRNA half-life in vitro by a protein capable of
binding to a coding region stability determinant
Genes Dev. 6 (4), 642-654 (1992)
JOURNAL 92217743
MEDLINE 1559612
PUBMED
REFERENCE 2 (bases 1 to 2223)
AUTHORS Herrick,D.J. and Ross,J.
TITLE The half-life of c-myc mRNA in growing and serum-stimulated cells:
influence of the coding and 3' untranslated regions and role of
ribosome translocation
Mol. Cell. Biol. 14 (3), 2119-2128 (1994)
JOURNAL 94158886
MEDLINE 8114742
PUBMED
REFERENCE 3 (bases 1 to 2223)
AUTHORS Prokipcak,R.D., Herrick,D.J. and Ross,J.
TITLE Purification and properties of a protein that binds to the
C-terminal coding region of human c-myc mRNA
J. Biol. Chem. 269 (12), 9261-9269 (1994)
JOURNAL 94179348
MEDLINE 8132663
PUBMED
REFERENCE 4 (bases 1 to 2223)
AUTHORS Leeds,P., Kren,H.T., Boylan,J.M., Betz,N.A., Steer,C.J.,
Gruppuso,P.A. and Ross,J.
TITLE Developmental regulation of CRD-BP, an RNA-binding protein that
stabilizes c-myc mRNA in vitro
Oncogene 14 (11), 1279-1286 (1997)
JOURNAL 97322234
MEDLINE 9178888
PUBMED
REFERENCE 5 (bases 1 to 2223)
AUTHORS Doyle,G.A., Betz,N.A., Leeds,P.F., Fleisig,A.J., Prokipcak,R.D. and
Ross,J.
TITLE The c-myc coding region determinant-binding protein: a member of a
family of KH domain RNA-binding proteins
Nucleic Acids Res. 26 (22), 5036-5044 (1998)
JOURNAL 98030504
MEDLINE 9801297
PUBMED
REFERENCE 6 (bases 1 to 2223)
AUTHORS Ross,J., Prokipcak,R.D., Leeds,P., Doyle,G.A.R., Betz,N.A. and
Fleisig,A.J.
TITLE Direct Submission
JOURNAL Submitted (25-APR-1998) McArdle Laboratory, University of Wisconsin
- Madison, 1400 University Ave., Madison, WI 53706, USA
FEATURES
source
1..2223
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
131..1864
/function="putative c-myc mRNA binding protein"
/notes="CRD-BP"
/product="coding region determinant binding protein"
/protein_id="AAC72743.1"
/db_xref="GI:3273749"
/translation="MNKLIIGNLNVTPADLEKVPAEHKISYSGQFLVKSQVAFVDC
PUEHWAKAIEITSGVEIQGRLEIEHSVPKKQSRKIOIRNIPQLRWEVLDSLLA
QYGVENCQVNTSEYAVNVTVSNREQTQAIKMLNGHLENHAKVSYIPDEQIT
QCFPENGRRGSGRQGPQGVPAAGAPAKQCPVDIPRLVPTQVGAIIKGEKATI
RNITKOTQSKIDVHRKENAGAAKAIKSHVSTPEGSSACKMILEIMHKAQDTKTADE
VPLKILAHNNFVRLIGKGRNLKVEQDTETKITISSIQDLTLNPNRTITVKGALIE
NCCRAEOETMKVREAYENDVAMSLQSHLIPGLNLAAVGLTPASSAVPPPSVTG
ADPYSFMOAPQSEWQVFIQAQVGAIIKGRGHKQLSRFASAKIAPPETIDSK
VRMVTITGPPEAOFKAQGRYKLEENFFGPKREVLETHIRVPASAAAGRVIGGK
TVNELQNLTAEEVVVPRDQTPDENQVIVKIIGHFYASQMAQRKIRDILAQVQKQHK
GQSNLAQARRK"
ORIGIN
Query Match 58.1%; Score 1162.6; DB 10; Length 2223;
Best Local Similarity 83.1%; Pred. No. 7.2e-282;
Matches 1416; Conservative 0; Mismatches 254; Indels 34; Gaps 7;
QY 30 CCGGGGAGGCATCATGAAGCTGAATGGCCACAGTTGGAGAACCATGCGCTGAAGGTCT 89
Db 525 CCAGGCAAGCTATCATGAAGCTAATGGCCATCACTGGAGAACCATGCGCTGAAGGTCT 584
QY 90 CCTACATCCCGATGAGCAGATAGCAGGAGCCTGAGATGGGCCCGCAGGGGGCTTTG 149
Db 585 CCTACATACCTGATGAGCAGATAACCAAGGCTCTGAGATGGGCGCTCTGGAGGCTTTG 644
QY 150 GCTCTCGGGTCAAGCCCGCCAGGGGTCACTGTGCGAGCGGGGGCCCGCAGCAACAGC 209
```

|      |   |      |
|------|---|------|
| 645  | GGTCTCGGGGCGACCCCGGCAAGGGGTGCGCCGTGGCAGCAGGGGCTTCACGCAAGCAGC    | 704  |
| 210  | AGCAAGTGGACATCCCCCTTCGGCTCCTGGTGCCACCCACGATATGTGGGTGCCATTATTG   | 269  |
| 705  | AGCCAGTGGACATCCCTCTCGGCTCCTGGTGCCCTACGCAGTATGTAGGCGCTATCATTTG   | 764  |
| 270  | GCARAGAGGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCCAGATAGACGTGC      | 329  |
| 765  | GCAAGAGGGGTGCCACATCCGAAACATCAAAAACAGACGCACTCCAAAATAGACGTGC      | 824  |
| 330  | ATAGGAAGGAGAACGCAAGTGCAGCTCAAAAAGCCATCAGTGTGCATCTCCAACCCCTCAGG  | 389  |
| 825  | ATAGGAAGGAGATCGGGCGCTGGGAGAGGCCATCAGCGTGCACTTCAACCCCTGAAG       | 884  |
| 390  | GCTGCTCCTCGGTTGTAAGATGATCTTGGAGATATATGCATAAAGAGGCTTAAGACACCA    | 449  |
| 885  | GCTGCTCCTCGCGTGCAAGATGATCTTGGAGATATATGCACAAGAGAGGCCAAAGACACCA   | 944  |
| 450  | AAACGGCTGACGAGGTTCCCTGAGAGTCTCTGGCCCATATACTTTGTAGGGGCTCTCA      | 509  |
| 945  | AAACGGCAGATGAAAGTTCCCTGAAGATCCTGGCTCATACACTTCTGTCGGGCGACTCA     | 1004 |
| 510  | TTTGCAAGGAGGACGGAACTCGAAGAGGTAGAGCAAGTACCGAGACAAAATCACCA        | 569  |
| 1005 | TTTGCAAGGAGGCGCGAACTCGAAGAGGTGGAGCAGGACACAGAGACGAAGATCACCA      | 1064 |
| 570  | TCCTCTGTTGCAAGACTTACCTTTACAACCTGAGAGACCATTCACTGTGAGAGGGG        | 629  |
| 1065 | TCTCATCGCTCCAGGACCTCAGCTCTATAACCTGAGAGACCATTCACTGTGAGAGGGG      | 1124 |
| 630  | CCATCGAGAAATGTTGCAAGGCCAGCAGAGAAATAATCAAGAAAGTTTCGGAGGCGCTATG   | 689  |
| 1125 | CCATTGAGAACTGTTGCAAGGCCAGCAGAGATCATGAGAAAGTTTCGAGAGGCTTACG      | 1184 |
| 690  | AGAAATGATGGCTGGCAATGAGC-----TCTCACTGATCCTCTGGGCTGAACTGGCTG      | 743  |
| 1185 | AGAAAGCAGCTGGCGCCATGAGCTTGCAAGTCCCACTCATCCTCTGGGCTTAACTGGCTG    | 1244 |
| 744  | CTGTAGGCTCTTTCCCAAGCTTCATCCAGCGCAGTCCCGCCGCTCCAGCAGCGTTACTG     | 803  |
| 1245 | CTGTAGGCTCTTCCCAAGCTTCATCCAGCGCTGTCCTCTCTCCAGCAGGTGCTACTG       | 1304 |
| 804  | GGGCTGCTCCTATAGCTCCTTTATGAGGCTCCCGAGCAGGAGATGGTGCAGTGTTTA       | 863  |
| 1305 | GGGCTGCTCCTATAGCTCCTTCATGAGGCTCCGAGCAGGAGATGGTACAAGTGTTC        | 1364 |
| 864  | TCCCGGCCAGGAGTGGGCGCCATCATCGCAGAGAGGGGCGACATCAACAGCTCT          | 923  |
| 1365 | TCCCGGCCAGGCTGTGGGCGCCATCATTGGCAAGAGGGGCGACATCAACACTCT          | 1424 |
| 924  | CCCGGTTTGCAGCGCCTCCATCAAGATTTGCACACCCGAAACCTGACTCCAAAGTTC       | 983  |
| 1425 | CCCGTTTCGCCAGCGCTCCATCAAGATTTGTCACACAGAAACCTGACTCCAAAGTTC       | 1484 |
| 984  | GTATGGTTATCATCACTGGACCGCCAGAGGCCAATTCAGGGCTCAGGAGAGATCTATG      | 1043 |
| 1485 | GAATGGTGTGTCATCACTGGACCCCGCGCTCATCGGCAAGAGGGCGAAACCGTGAATG      | 1544 |
| 1044 | GCAAACTCAAGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTCAAGCTGGAGACCCACA     | 1103 |
| 1545 | GCAAACTCAAGAGAGAAATTTCTTTGGTCCCAAGGAGGAAGTAAAGCTAGAGACCCACA     | 1604 |
| 1104 | TACGTGTGCCAGCATCAGCAGCTGGCGGGTCAATTGGCAAGGTGGAAAAACCGTGAACG     | 1163 |
| 1505 | TACGGGTTCCGGCTTTCAGCAGCCCGCGCGTCACTCGGCAAGAGGGCGGCAAAACCGTGAATG | 1664 |
| 1164 | AGTTGCAGAAATTTCAACGCGCAGCTCAGGTGGTGAATACCAAGAGACACAGCCCTCATGAGA | 1223 |
| 1565 | AGCTGCAGAACTTGACTGTGAGCTGAGGTGATGATGCCAGAGACAGACCCCGATGAGA      | 1724 |
| 1224 | ACGACCAAGGTCATCGTGAATAATCATCGCAATTTCTATGCCAGTTCAGATGGCTCAACGGA  | 1283 |
| 1725 | ACGACCAAGTCATTTGTTAAGATCATCGCAATTTTATGCAAGCCAGATGGCTCAGCGGA     | 1784 |

|            |  |  |                 |
|------------|--|--|-----------------|
| QY         | 1284   | AGATCCGAGACATCTGGCCCGCGGTTAAGCAGCAGCATCAGAGGGGACAGAGTAACCAAG   | 1343            |
| Db         | 1785   | AGATCCGAGACATCTGGCTCAAGTTAAGCAACAGCACACAGAGGACAGAGCAACCTCG     | 1844            |
| QY         | 1344   | CCAGGCGACGAGGAGTGA-CCAGGCCCTCCCTGTGCCCTTNGAGTCCAGGACAAACAACG   | 1402            |
| Db         | 1845   | CCAGGTCGGAGGAATGACCCGCCCTCTGTCCCATGCTGCTCCAGATCATGCGAGG        | 1904            |
| QY         | 1403   | GGCAGAA-----ATCGAGAGTGTCTCTCCCGCGCAGCCTGAG                     | 1441            |
| Db         | 1905   | AGGAACACAGAACTGGAGGGGGGGTGGAGGGCCGGTGTGTTTTTCCAGCAGGCTGAG      | 1964            |
| QY         | 1442   | AATGAGTGGGAATCCGGGACACNCTGGCCGGGCTGTAGATCAGGTTTGCCACTTGATTG    | 1501            |
| Db         | 1965   | AATGAGTGGGAATCAGGG-CATTGGGCGCTGGGTGGAGATCAGGTTTGCACACTGTAATG   | 2023            |
| QY         | 1502   | AGAAAGATGTTCCAGTGAAGCAACCTGATCTNTCAGGCCCAACACACCCACCCCAATTGCGC | 1561            |
| Db         | 2024   | AGACAATGTTCCAGTGAAGATCCCTGATCTCTCGCCCCCAA--TTGAGCCAGCTGSCCA    | 2081            |
| QY         | 1562   | CACACTGTNTGCCCTCGGGGTGCAGAAATTTNAGCGCAGGCACCTTTTAAACGTGGA      | 1621            |
| Db         | 2082   | CAGCCACCCCTTGGAAATATCACCATTCGAATCATAGCTTGGGTGTGCTTTTAAACGTGGA  | 2141            |
| QY         | 1622   | TTGTTTAAAGAAAGCTCTCCAGGCGCCACCAAGAGGGTGAATCACACTCAGTGGGAAGAA   | 1681            |
| Db         | 2142   | TTGTCT--TGAAGTTCTTCAGCTCCATGGA-AGGATGGTCAATCCCAATGGGGNAGA      | 2198            |
| QY         | 1682   | AAATAAAATTTCCCTCAGGTTTTA                                       | 1705            |
| Db         | 2199   | GAATAAAATTTCCCTCAGGTTTTA                                       | 2222            |
| RESULT 13  |  |  |                 |
| BC051679   |  | 2444 bp  | linear          |
| LOCUS      |  |  |                 |
| DEFINITION | Mus musculus insulin-like growth factor 2, binding protein 1, mRNA   |  | ROD 08-OCT-2003 |
|            | (CDNA clone MGC:60613 IMAGE:30008106), complete cds.   |  |                 |
| ACCESSION  | BC051679.1   | GI:30354043  |                 |
| VERSION    | MGC.   |  |                 |
| KEYWORDS   | Mus musculus (house mouse)   |  |                 |
| SOURCE     | Mus musculus   |  |                 |
| ORGANISM   | Mus musculus   |  |                 |
| REFERENCE  |  |  |                 |
| AUTHORS    | Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altshul,S.F., Zebberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Sapotnick,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Utsin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullay,J.S., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaxathne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahy,J., Helton,E., Kettaman,M., Madan,A., C., Shevchenko,Y., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalski,U., Smalls,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A. |  |                 |
| TITLE      | Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences   |  |                 |
| JOURNAL    | Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  |  |                 |
| MEDLINE    | 22388257   |  |                 |
| PubMed     | 12477932   |  |                 |
| REFERENCE  | 2 (bases 1 to 2444)  |  |                 |
| AUTHORS    | Strausberg,R.  |  |                 |
| TITLE      | Direct Submission  |  |                 |



**JOURNAL**

REMARK  
COMMENT

Submitted (30-APR-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: Dr. Kenneth R. Boheler (National Institute on Aging, USA)  
cDNA Library Preparation: Yulan Piao and Minoru Ko (National Institute on Aging, NIH: <http://lgsun.grc.nia.nih.gov/cDNA/>)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (fNLN)  
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
Gunnaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAK Plate: 111 Row: e Column: 24  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6753517.  
Location/Qualifiers

## FEATURES

source

```
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129/Sv x 129/Sv-CP"
/db_xref="taxon:10090"
/clone="MGC:60613 IMAGE:30008106"
/tissue_type="ES Cell, undifferentiated, mouse, R1
Embryonic stem cells"
/clone_lib="NIA Mouse Undifferentiated ES Cell cDNA
Library (long)"
/lab_host="DH10B"
```

gene

```

1. 2432
/gene="Igf2bp1"
/notes="synonyms: IMP-1, Neilsen, CRD-3P"
/db_xref="LocusID:140486"
/db_xref="MGI:1890357"
2957.2028

```

505

```

/codon_start=1
/product="insulin-like growth factor 2, binding protein 1"
/product_id="AAHS1679.1"
/db_xref="GI:3035404"
/db_xref="LocusID:140486"
/translation="MNKLYIGNESVTPADLKVEAEHKISTSGYFLKSGYAFVDC
PDSHMAKAIETFGSGVQGGKRLSEIHSPVQKRSKQIQRINPQLRVEYLDLSLA
QYGVCEQVQNTSEPTAVNTVYSNRQETQAINKLNGHLENHALKWSYIPDEQIT
QYGVCEQNGRGSGRQSGPAGAPQAPQVDFILRLPLVITVYGAALIGKEGATI
RNITKQTSKIDVHRKNGAAGAAKAI SVHSTPEGCSACMKMLIEHMEKAEKGTADG
VPLKLIAHNFVVRGIRKEGNRLKVEQDTEKTIISLQDLTYNPERITTVKGAIE
NCCRAQETMKKVRAYENEDVAANSLSQLHLPGLNLAAGLTFPASSAVPPPSVTVG
AAYTSFQMAQEQMVQVPIPAQVGAALIGKGGHQLQLGRFSASIKAPPTPDGK
VRMVVITGPPEAQVGAQGRIPQAQGLKEENFFGSGKEVLETHIVPASAAGRVITGSGK
TVNELNQLAAEVVTPDQTPENDQIVKLIIGHFYSASQMAQRKIDPILAQVQQRQK
GOSNLQAQAEK"

```

misc feature

```

350: 730
/note="RRM; Region: RNA recognition motif"
/db_xref="CDD:smart00360"
877: 1086

```

misc feature

```

077: 1000
/note="KH; Region: X homology RNA-binding domain"
/db xref="CDD:smart00322"

```

misc\_feature

```

/note="KH; Region: K homology RNA-binding domain"
/db xref="CDD:smart00322"

```

## ORIGIN

### Query Match

68.0%; Score 1161.6; DB 10; Length 2444;

[illegible]

|            |  |  |        |
|------------|--|--|--------|
| Db         | 1709   | GCAAACTAAAGAGAGAAATTTCTTTGGTCCCAAGGAGAGTAAGTAAGCTAGAGACCCACA | 1768   |
| Qy         | 1104   | TACGTGTCCAGCATCAGCAGCTGGCGGGTCAATGGCAAGGTGGAAAAACGGTGAACG    | 1163   |
| Db         | 1769   | TACGGTTCGGCTTCAGCAGCGCGGTGTCTATCGCAAGGCGCAAAACGGTGAATG       | 1828   |
| Qy         | 1164   | AGTTGCAGAAATTCAGCGCAGCTGAGTGGTGTACCAAGAGACCAACCCCTGTATGAGA   | 1223   |
| Db         | 1829   | AGCTGCAGAACTTGACCGCAGCTGAGTGGTGTAGTCCCAAGAGACCAACCCCGATGAGA  | 1888   |
| Qy         | 1224   | ACGACCGCTCATCTGTGAATATATCGACATTTCTATGCCAGTCACATGGCTCAACGGA   | 1283   |
| Db         | 1889   | ACGACCAAGTCAATTTGAAGATATCGGACATTTCTATGCCAGTGGCTCAGCGGA       | 1948   |
| Qy         | 1284   | AGATCCGAGACATCTGGCCCGAGTTAAGCAGCAGCATCAGAGGACAGAGTAACAGG     | 1343   |
| Db         | 1949   | AGATCCGAGACATCTGGCTCAAGTTAAGCAACAGCACCAAGAGGACAGAGCAACCTGG   | 2008   |
| Qy         | 1344   | CCCAGGACGGAGAGTGA--CCAGCCCTCCCTGTCCCTTNGAGTCCAGGACAAACAG     | 1402   |
| Db         | 2009   | CCCAGGCTCGAGGAAGTGAACCCCGCCCTCTCTGTCCATTTGGCTCAGATCAGCAGG    | 2068   |
| Qy         | 1403   | GGCGAA-----ATCGAGAGTGTGTCTCTCCCGGCGAGCCCTGAG                 | 1441   |
| Db         | 2069   | AGGAACACAGAACTGGAGGGCGGTGGAGGGCGGTGTGTCTCTCCAGCAGGCGCTGAG    | 2128   |
| Qy         | 1442   | AATGATGGGAATCCGGACACATNTGGCGGGCTGTAGATCAGATTTTCCCACTTGATTG   | 1501   |
| Db         | 2129   | AATGATGGGAATCAGGG--CATTTGGGCTCGCTGGAGATCAGTGTTCGACACTGTCTTG  | 2187   |
| Qy         | 1502   | AGAAAGATGTTCCAGTGAAGAACCTGATCTNTCAGCCCCCAACACCCACCAATTTGGCC  | 1561   |
| Db         | 2188   | AGAACATGTTCCAGTGAAGATCTGATCTCTCGCCCCCA--TTGAGCCAGCTGGCCA     | 2245   |
| Qy         | 1562   | CAACACTGNTGCCCTCGGGGTGTGAAATTTAGGCGCAAGCATTCTTAAACGTGGA      | 1621   |
| Db         | 2246   | CAGCCACCCCTTGGAAATATCACCATTGCAATCATAGCTTGGGTGTCTTAAACGTGGA   | 2305   |
| Qy         | 1622   | TGTTTAAAGAGCTCTCAGGCCCCCAAGAGGGTGGATCACACCTCAGTGGGAAGAA      | 1681   |
| Db         | 2306   | TGTGCT--TGAAGTCTCCAGCTCCATGGAAGATGGGTGATCCAGTCCAGTGGGAAGAG   | 2363   |
| Qy         | 1682   | AAATAAAATTTCTTCAGGTTTTAAA                                    | 1708   |
| Db         | 2364   | AAATAAAATTTCTTCAGGTTTTTAA                                    | 2390   |
| RESULT 14  |  |  |        |
| AF541940   |  |  |        |
| LOCUS      |  |  |        |
| DEFINITION |  |  |        |
| ACCESSION  | AF541940   |  |        |
| VERSION    | AF541940.1   | GI:27464837  |        |
| KEYWORDS   |  |  |        |
| SOURCE     | Rattus norvegicus (Norway rat)   |  |        |
| ORGANISM   | Rattus norvegicus  |  |        |
| REFERENCE  |  |  |        |
| AUTHORS    | Eom, T., Antan, J.N., Singer, R.H. and Bassell, G.J.   |  |        |
| TITLE      | Localization of a beta-actin messenger ribonucleoprotein complex with zipcode-binding protein modulates the density of dendritic filopodia and filopodial synapses |  |        |
| JOURNAL    | J. Neurosci. 23 (32), 10433-10444  |  | (2003) |
| MEDLINE    | 22979172   |  |        |
| PUBMED     | 14614102   |  |        |
| REFERENCE  |  |  |        |
| AUTHORS    |  |  |        |
| TITLE      |  |  |        |
| JOURNAL    |  |  |        |
| MEDLINE    |  |  |        |
| PUBMED     |  |  |        |
| REFERENCE  |  |  |        |
| AUTHORS    | Eom, T., Singer, R.H. and Bassell, G.J.  |  |        |
| TITLE      | Direct Submission  |  |        |
| JOURNAL    | Submitted (27-AUG-2002) Neuroscience, Albert Einstein College of Medicine, 1300 Morris Park Ave., Bronx, NY 10461, USA   |  |        |

| FEATURES | source | Location/Qualifiers   |
|----------|--------|---|
| CDS      |        | <p>1..2105</p> <p>/organism="Rattus norvegicus"</p> <p>/mol_type="mRNA"</p> <p>/strain="Sprague-Dawley"</p> <p>/db_xref="taxon:10116"</p> <p>223_..1956</p> <p>/note="rZBP1"</p> <p>/codon_start=1</p> <p>/product="b-actin zipcode binding protein 1"</p> <p>/protein_id="AA016210.1"</p> <p>/db_xref="GI:27464838"</p> <p>/translation="MNKLYIGNLNESTVPADLEKVFPAEHKISYSGQFLVKSQYAFVDC<br/>PDEHWAKEIETFSGLVQLQKLEIEHSVTKORSKIQIRNIIPQLRWEVLDSLLA<br/>QYGTVECEQNTSETAIVNVTYSNREOTQAIKLNHOLENHALKVSYPDEQIAI<br/>QAPENRKGFGSGRQOPGQPSVAGAPAKQQQVDIPLRLVPTQVYGAIIKKEGAT<br/>RNITQTSKIDVHRKENAGAAEKASIVHSITPEGSSACAMLEIMKEAKDTKTADE<br/>VPLKTLAHNNFVRLIGREGRNKKVQDTEKITISLQDLTLNPERITTVKALG<br/>NCCRAEQIMKKVREAYENDVAMSLQSHLIPGLNLAAGLFPASSAVPPPPSVTG<br/>AAPYGSFQAPQEMVQVFIQAQVAGNFIKGGQHIKQLSRFASASIKIAPPETPSK<br/>VRMVITGPPBAQFKAQGRIVKGLKEENFGPKVEVKLETHIRVPSAAGRIVKGGK<br/>TYNELNLTAEEVVPRDTPDENDQIVIKLIIGHFYASQMAQRKIRDILAVKQKHQK<br/>QGSNQQAQRK"</p> |
| ORIGIN   |        | <p>Query Match 65.1%; Score 1112.4; DB 10; Length 2105;</p> <p>Best Local Similarity 86.2%; Pred. No. 3.2e-269;</p> <p>Mismatches 1285; Conservative 0; Mismatches 178; Indels 27; Gaps 4;</p>  |
| Qy       | 30     | CCCCGGGAGCCATCATGAGCTGAATGGCCACCATTTGGAGAACCATCCCTGAGGTCT 89  |
| Db       | 617    | CCAGGCAAGCCATCATGAGCTAAATGGCCATCACTGGAGAACCATCCCTGAGGTCT 676  |
| Qy       | 90     | CTTACATCCCCCATGAGCAGATAGCACAGGGACCTTGAGATCGGGCGGAGGGGCTTTG 149  |
| Db       | 677    | CTTACATCTCTATGAGCAGATAGCACAGGTCTCTGAGATCGGGCGGCTTGGAGGCTTTG 736   |
| Qy       | 150    | GTCTCGGGGTGAGCCCGCCAGGCTCACCTGTGGCAGGGGGGCCCCAGCAAGCAGC 209   |
| Db       | 737    | GGTCTCGAGGCGAGCCAGGCGGGGTCAACCGTAGCAGCAGGGGCGGCGAGCAAGCAGC 796  |
| Qy       | 210    | ACCAAGTGACATCCCCCTTCGGCTCTCGTGTGCCACCCAGTATGTGGTGTCATTATTG 269  |
| Db       | 797    | AGCAGGTGACATCCCTCTCGGCTCTGTGTGCCACCGAGTATGTAGGGCGCCATCATTTG 856   |
| Qy       | 270    | GCAAGGAGGGGGCCACCATCCGCAACATCAAAAAACAGCCAGTCCCAAGATAGACGTGC 329   |
| Db       | 857    | GCAAGGAGGGTGCACCATCCGAAACATCAAAAAACAGCTCAGTCCAAATAGACGTGC 916   |
| Qy       | 330    | ATAGGAGAGAGAACCGAGGTGCAGTGAAGAACCCATGCTGTGCTCCACCTCCCTGAGG 389  |
| Db       | 917    | ACAGGAGAGAGAACCGGGGGCTGCGGAGAGGCGCCATCAGTGTGCTCCAGCCCTCGAAG 976   |
| Qy       | 390    | GCTGCTCTCCGCTTGTGAATGATCTTTGAGATATATGCTAAAGAGGCTAAGGACACCA 449  |
| Db       | 977    | GCTGCTCTCCGCGCTTGAATGATCTTTGAGATATATGCAACAGGAGGCAAGGACACCA 1036   |
| Qy       | 450    | AAACGGCTGACAGGTTCCCTCTGAAGATCCTGGCCCAATAATACTTTGTAGGGCGCTCTCA 509   |
| Db       | 1037   | AGACGGCAGACGAAGTTCCCTCTGAAGATCTTGCTCATATAACTTCGTGGGGGACTCA 1096   |
| Qy       | 510    | TTGGCAAGGAGGACCGGAACCTTGAAGAGGTAGACGAAGATACCGAGACAAATAATCACCA 569   |
| Db       | 1097   | TTGGAAAGGAGGCGCGAACTTTGAAGAAAGTGGAGCAGGATACAGAGACTAAGATACCA 1156  |
| Qy       | 570    | TCTCTCTGTTTGAAGACCTTACCTTTTACACCTCGAGAGGACCATCACTGTGTGAAGGGGG 629   |
| Db       | 1157   | TCTCTCTCGTCCAGGATCTCACGCTCTATAACCTTGAAGAGGACCATCACTGTGTGAAGGGGG 1216  |
| Qy       | 630    | CCATCGAGAAATTTGTCAGGGCGGAGCAGAAATAATGAAGAAAGTTTCGGGAGGCGCTATG 689   |
| Db       | 1217   | CCATTGAGAACTGCTGAGGGCTGAGCAGAGATCATGAAGAAAGTTTCGAGAGGCTTAGC 1276  |





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 13, 2004, 20:24:00 ; Search time 792 Seconds  
(without alignments)  
10510.849 Million cell updates/sec

Title: US-09-270-437D-5

Perfect score: 1708

Sequence: 1 agggacgtcgccagccgc.....atttcctcagggttttaaaa 1708

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3183909 seqs, 2436941669 residues

Total number of hits satisfying chosen parameters: 6367818

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*

2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*

3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*

4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*

5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*

6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*

7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*

8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*

9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*

10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*

11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*

12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*

13: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*

14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*

15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*

16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*

17: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*

18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*

19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description        |
|------------|--------|-------------|--------|-------|--------------------|
| 1          | 1703   | 99.7        | 1708   | 9     | US-09-899-651-5    |
| 2          | 1669.2 | 97.7        | 1945   | 9     | US-09-899-651-7    |
| 3          | 1293.6 | 75.2        | 1738   | 16    | US-10-062-674-2082 |
| 4          | 1165   | 68.2        | 2224   | 9     | US-09-873-637-1    |
| 5          | 1085.4 | 63.5        | 1182   | 9     | US-09-764-864-327  |
| 6          | 1069.8 | 62.6        | 1129   | 9     | US-09-764-864-745  |
| 7          | 582.2  | 34.1        | 1740   | 9     | US-09-735-705-347  |
| 8          | 582.2  | 34.1        | 1740   | 9     | US-09-850-716A-347 |
| 9          | 582.2  | 34.1        | 1740   | 9     | US-09-837-778-347  |
| 10         | 582.2  | 34.1        | 1740   | 13    | US-10-007-700-347  |
| 11         | 582.2  | 34.1        | 1740   | 13    | US-10-117-982-347  |
| 12         | 582.2  | 34.1        | 1740   | 15    | US-10-117-982-478  |
| 13         | 582.2  | 34.1        | 1740   | 16    | US-10-313-986-347  |
| 14         | 582.2  | 34.1        | 1740   | 16    | US-10-313-986-478  |

15 582.2 34.1 1743 9 US-09-897-778-447 Sequence 447, App

16 582.2 34.1 1743 9 US-09-897-778-450 Sequence 450, App

17 582.2 34.1 1743 13 US-10-007-700-447 Sequence 447, App

18 582.2 34.1 1743 13 US-10-007-700-450 Sequence 450, App

19 582.2 34.1 1743 15 US-10-117-982-447 Sequence 447, App

20 582.2 34.1 1743 15 US-10-117-982-450 Sequence 450, App

21 582.2 34.1 1743 16 US-10-313-986-447 Sequence 447, App

22 582.2 34.1 1743 16 US-10-313-986-450 Sequence 450, App

23 582.2 34.1 1764 9 US-09-850-716A-428 Sequence 428, App

24 582.2 34.1 1764 9 US-09-897-778-428 Sequence 428, App

25 582.2 34.1 1764 13 US-10-007-700-428 Sequence 428, App

26 582.2 34.1 1764 15 US-10-117-982-428 Sequence 428, App

27 582.2 34.1 1764 16 US-10-313-986-428 Sequence 428, App

28 582.2 34.1 1799 16 US-10-313-986-485 Sequence 485, App

29 582.2 34.1 34555 15 US-10-117-982-479 Sequence 479, App

30 582.2 34.1 34555 16 US-10-313-986-479 Sequence 479, App

31 580.6 34.0 4159 9 US-09-899-651-4 Sequence 4, Appli

32 580.6 34.0 4181 9 US-09-735-705-175 Sequence 175, App

33 580.6 34.0 4181 9 US-09-954-456-715 Sequence 715, App

34 580.6 34.0 4181 9 US-09-850-716A-175 Sequence 175, App

35 580.6 34.0 4181 10 US-09-466-396A-175 Sequence 175, App

36 580.6 34.0 4181 13 US-10-007-700-175 Sequence 175, App

37 580.6 34.0 4181 15 US-10-117-982-175 Sequence 175, App

38 580.6 34.0 4181 16 US-10-313-986-175 Sequence 175, App

39 580.6 34.0 4181 16 US-10-116-802-145 Sequence 145, App

40 580.6 34.0 4334 13 US-10-116-802-145 Sequence 145, App

41 572.6 33.5 1740 15 US-10-117-986-483 Sequence 483, App

42 572.6 33.5 3283 9 US-09-899-651-8 Sequence 8, Appli

43 504 29.5 3667 15 US-10-097-340-146 Sequence 146, App

44 504 29.5 3667 17 US-10-648-593-48 Sequence 48, Appli

45 504 29.5 3667 17 US-10-648-593-48 Sequence 48, Appli

#### ALIGNMENTS

##### RESULT 1

US-09-899-651-5  
; Sequence 5 Application US/09899651  
; Patent No. US20020111470A1  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Yao-Tseng  
; APPLICANT: Gure, Ali  
; APPLICANT: Tsang, Solam  
; APPLICANT: Stockert, Elisabeth  
; APPLICANT: Jager, Elke  
; APPLICANT: Knuth, Alexander  
; APPLICANT: Old, Lloyd J.  
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer  
; TITLE OF INVENTION: Associated  
; TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof  
; FILE REFERENCE: LUD 5538  
; CURRENT APPLICATION NUMBER: US/09/899,651  
; CURRENT FILING DATE: 2001-07-06  
; PRIOR APPLICATION NUMBER: US/09/061,709  
; PRIOR FILING DATE: 1998-04-17  
; NUMBER OF SEQ ID NOS: 8  
; SEQ ID NO 5  
; LENGTH: 1708  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
US-09-899-651-5

Query Match 99.7%; Score 1703; DB 9; Length 1708;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AGGACGCTGCGCCAGCCCGCAGTTTACCCGGGAGCCCATCATGAAGCTGAATGGCCA 60  
Db 1 AGGACGCTGCGCCAGCCCGCAGTTTACCCGGGAGCCCATCATGAAGCTGAATGGCCA 60  
QY 61 CCAGTTGGAGACCATGCGCCTGAAGGTCTCTACATCCCGATGAGCAGATGACACAGG 120

Db 61 CCAGTTGGAGAACATGCCCTGAAGGCTCTCTACATCCCGATGAGCAGATACACAGG 120  
Qy 121 ACCTGAGAAATGGCGCGAGGGGCTTTGGCTTCGCGGGTACGCCCGCCAGGGCTCAC 180  
Db 121 ACCTGAGAAATGGCGCGAGGGGCTTTGGCTTCGCGGGTACGCCCGCCAGGGCTCAC 180  
Qy 181 TGTGCGAGCGGGGCGCCAGCAGAGCAGCAGTGGATCATCCCTCTCGGCTCTCTGGT 240  
Db 181 TGTGCGAGCGGGGCGCCAGCAGAGCAGCAGTGGATCATCCCTCTCGGCTCTCTGGT 240  
Qy 241 GCCCACCAGTATGTGGTGCATTATTTGGCAAGAGAGGGGGCCACATCCGCAACATCAC 300  
Db 241 GCCCACCAGTATGTGGTGCATTATTTGGCAAGAGAGGGGGCCACATCCGCAACATCAC 300  
Qy 301 AAAACAGACCCAGTCCAGATAGACGTGATAGAGAGAGGAGGAGGAGGAGGAGGAGG 360  
Db 301 AAAACAGACCCAGTCCAGATAGACGTGATAGAGAGAGGAGGAGGAGGAGGAGGAGG 360  
Qy 361 AGCCATCAGTGTGCACTCCACCCCTGAGGGCTGCTCTCCCTGTTGTAAGATGATCTTGA 420  
Db 361 AGCCATCAGTGTGCACTCCACCCCTGAGGGCTGCTCTCCCTGTTGTAAGATGATCTTGA 420  
Qy 421 GATTATGATTAAGAGGCTAAGACACCAAAACGGCTGAGAGGTTCCCTGAGATCCT 480  
Db 421 GATTATGATTAAGAGGCTAAGACACCAAAACGGCTGAGAGGTTCCCTGAGATCCT 480  
Qy 481 GGCCCATTAATCTTTGTAGGGCTCTCATTTGGCAAGAGGAGGAGGAGGAGGAGGAGG 540  
Db 481 GGCCCATTAATCTTTGTAGGGCTCTCATTTGGCAAGAGGAGGAGGAGGAGGAGGAGG 540  
Qy 541 AGAGCAAGATACGAGACAAAATCAACATCTCTCTGTTGCAAGACCTTACCTTTACAA 600  
Db 541 AGAGCAAGATACGAGACAAAATCAACATCTCTCTGTTGCAAGACCTTACCTTTACAA 600  
Qy 601 CCCTGAGAGGACCATCACTGTGAAGGGGCGCATCGAAGATTGTTGAGGGCCGAGCAGGA 660  
Db 601 CCCTGAGAGGACCATCACTGTGAAGGGGCGCATCGAAGATTGTTGAGGGCCGAGCAGGA 660  
Qy 661 AATAATGAAGAAAGTTCGGAGGCTATGAGAAATGATGTGGCTGCCATGAGCTCTACCT 720  
Db 661 AATAATGAAGAAAGTTCGGAGGCTATGAGAAATGATGTGGCTGCCATGAGCTCTACCT 720  
Qy 721 GATCCCTGGCTGAACCTGGCTGTAGTCTTTCCAGCTTCATCCAGCGCAGTCCC 780  
Db 721 GATCCCTGGCTGAACCTGGCTGTAGTCTTTCCAGCTTCATCCAGCGCAGTCCC 780  
Qy 781 GCGGCTCCCGAGCAGCTTACTGGGGCTGCTCCCTATAGTCTCTTTATGAGGCTCCCGA 840  
Db 781 GCGGCTCCCGAGCAGCTTACTGGGGCTGCTCCCTATAGTCTCTTTATGAGGCTCCCGA 840  
Qy 841 CGAGAGATGTGAGGTTTATCCCGCCAGCAGTGGCGGCGCATCATCGCGAGAA 900  
Db 841 CGAGAGATGTGAGGTTTATCCCGCCAGCAGTGGCGGCGCATCATCGCGAGAA 900  
Qy 901 GGGGAGACATCAACAGCTCTCCGGTTTGCAGGCGCTCCATCAAGATTGACCAACC 960  
Db 901 GGGGAGACATCAACAGCTCTCCGGTTTGCAGGCGCTCCATCAAGATTGACCAACC 960  
Qy 961 CGAAACATGATCTCAAGTTGCTATGTTTATCATCATGAGCGGCGAGGGCCCAATT 1020  
Db 961 CGAAACATGATCTCAAGTTGCTATGTTTATCATCATGAGCGGCGAGGGCCCAATT 1020  
Qy 1021 CAAGCTCAGGAGAACTATGCAAACTCAAGAGGAGAACTTCTTGGTCCCAAGGA 1080  
Db 1021 CAAGCTCAGGAGAACTATGCAAACTCAAGAGGAGAACTTCTTGGTCCCAAGGA 1080  
Qy 1081 GGAAGTGAAGTGGAGACCAATACGTGTGCGAGCATCAGCAGTGGCGGGTCAATGG 1140  
Db 1081 GGAAGTGAAGTGGAGACCAATACGTGTGCGAGCATCAGCAGTGGCGGGTCAATGG 1140  
Qy 1141 CAAGGTGGAAGACGTTGACGATTCGAAATTTGACGACGCTGAGGTGGTAGTACC 1200

Db 1141 CAAGGTGGAAGAAACGTTGAACGAGTTGACAGATTTGACGCGAGCTGAGGTGGTAGTACC 1200  
Qy 1201 AAGAGACAGACCCCTGATGAGAACGACGACAGGTCACTGTAATAATCATCGACATTTCTA 1260  
Db 1201 AAGAGACAGACCCCTGATGAGAACGACGACAGGTCACTGTAATAATCATCGACATTTCTA 1260  
Qy 1261 TGCCAGTCAGATGGCTCAACGGAAGATCCGAGAATCTGCGCCCGAGGTTAAGCAGCAGCA 1320  
Db 1261 TGCCAGTCAGATGGCTCAACGGAAGATCCGAGAATCTGCGCCCGAGGTTAAGCAGCAGCA 1320  
Qy 1321 TCAGAGGAGCAGATGACAGGCGCCAGGCGACGAGGAGTACAGCCCTCCCTGTTCC 1380  
Db 1321 TCAGAGGAGCAGATGACAGGCGCCAGGCGACGAGGAGTACAGCCCTCCCTGTTCC 1380  
Qy 1381 CTTNAGTCCAGGACCAACACGCGGCAGAAATTCGAGAGTGTCTCTCCCGCAGGCGCTGA 1440  
Db 1381 CTTNAGTCCAGGACCAACACGCGGCAGAAATTCGAGAGTGTCTCTCCCGCAGGCGCTGA 1440  
Qy 1441 GAATGAGTGGAAATCCGGGACACNTGGCCGGGCTGTAGATCAGGTTTGCCCACTTGATT 1500  
Db 1441 GAATGAGTGGAAATCCGGGACACNTGGCCGGGCTGTAGATCAGGTTTGCCCACTTGATT 1500  
Qy 1501 GAGAAAGATGTTCCAGTCAGGAAACCTGATCTNTCAGCCCCCAACACACCCCAATTGGC 1560  
Db 1501 GAGAAAGATGTTCCAGTCAGGAAACCTGATCTNTCAGCCCCCAACACACCCCAATTGGC 1560  
Qy 1561 CCACACCTGTTNCCCTCGGGGTGTGAGAAATTTAGCGCAAGGCACCTTTTAAACGTGG 1620  
Db 1561 CCACACCTGTTNCCCTCGGGGTGTGAGAAATTTAGCGCAAGGCACCTTTTAAACGTGG 1620  
Qy 1621 ATTGTTTAAAGAGCTCTCCAGGCGCCACCAAGAGGTGGATCACCTCAGTGGGAAGA 1680  
Db 1621 ATTGTTTAAAGAGCTCTCCAGGCGCCACCAAGAGGTGGATCACCTCAGTGGGAAGA 1680  
Qy 1681 AAAATAAAATTTCTTCAGGTTTAAAA 1708  
Db 1681 AAAATAAAATTTCTTCAGGTTTAAAA 1708

RESULT 2

US-09-899-651-7  
; Sequence 7, Application US/09899651  
; Patent No. US20020111470A1  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Yao-Tsung  
; APPLICANT: Cure, Ali  
; APPLICANT: Tsang, Solam  
; APPLICANT: Stockert, Elisabeth  
; APPLICANT: Jager, Elke  
; APPLICANT: Knuth, Alexander  
; APPLICANT: Old, Lloyd J.  
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer  
; TITLE OF INVENTION: Associated  
; TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof  
; FILE REFERENCE: LUD 5538  
; CURRENT APPLICATION NUMBER: US/09/899,651  
; CURRENT FILING DATE: 2001-07-06  
; PRIOR APPLICATION NUMBER: US/09/061,709  
; PRIOR FILING DATE: 1998-04-17  
; NUMBER OF SEQ ID NOS: 8  
; SEQ ID NO 7  
; LENGTH: 1946  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
US-09-899-651-7

Query Match 97.7%; Score 1669.2; DB 9; Length 1946;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1676; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
Qy 30 CCGGGGAGCCCATCATGAGCTGAATGGCCACAGTGGAGAACCATGCCCTGAAGGTCT 89

Db 268 CCAGCAAGCCATCATGAAGCTGAATGGCCACCAAGTTGGAGAACCATGTCCTCGAAGGTCT 327  
QY 90 CTTACATCCCGATGAGCAGATAGCACAGGAGCTTGAGAAATGGGCGCGAGGGGCTTTG 149  
Db 328 CTTACATCCCGATGAGCAGATAGCACAGGAGCTTGAGAAATGGGCGCGAGGGGCTTTG 387  
QY 150 GCTCTCGGGGTAGCCCCCGCAGGGCTCACTGTGGCAGCGGGGGCCCCAGCCAGCAGC 209  
Db 388 GCTCTCGGGGTAGCCCCCGCAGGGCTCACTGTGGCAGCGGGGGCCCCAGCCAGCAGC 447  
QY 210 AGCAAGTGGACATCCCTCGCTTGGCTCTGTGGTGGCCACCCAGTATGTGGTGGCTATTTG 269  
Db 448 AGCAAGTGGACATCCCTCGCTTGGCTCTGTGGTGGCCACCCAGTATGTGGTGGCTATTTG 507  
QY 270 GCAAGAGGGGGCCACCATCCGCAACATCAAAAACAGACCCAGTCCAAAGATAGACGTGC 329  
Db 508 GCAAGAGGGGGCCACCATCCGCAACATCAAAAACAGACCCAGTCCAAAGATAGACGTGC 567  
QY 330 ATAGGAAGAGAACGACAGGTGAGCTGAAAAGCCATCAGTGTGCACTCCACCCCTGAGG 389  
Db 568 ATAGGAAGAGAACGACAGGTGAGCTGAAAAGCCATCAGTGTGCACTCCACCCCTGAGG 627  
QY 390 GCTGCTCTCGCTTGTAAAGTATCTTTGGAGATTATGCATAAAGAGGGCTAAGGACACCA 449  
Db 628 GCTGCTCTCGCTTGTAAAGTATCTTTGGAGATTATGCATAAAGAGGGCTAAGGACACCA 687  
QY 450 AAAAGGCTGACAGGTTCCCTGAGATCTGGGCCATPAATACTTTGTAGGGGCTCTCA 509  
Db 688 AAAAGGCTGACAGGTTCCCTGAGATCTGGGCCATPAATACTTTGTAGGGGCTCTCA 747  
QY 510 TTGGCAAGGAAGGAGCGAACTGAGAGAGGTAGAGCAAGTACCGAGACAAATACCA 569  
Db 748 TTGGCAAGGAAGGAGCGAACTGAGAGAGGTAGAGCAAGTACCGAGACAAATACCA 807  
QY 570 TCTCCTCTGCAAGACCTTACCCTTTACACCTCTGAGAGGACCATCACTGTGAGGGGG 629  
Db 808 TCTCCTCTGCAAGACCTTACCCTTTACACCTCTGAGAGGACCATCACTGTGAGGGGG 867  
QY 630 CCATCGAAGTGTGAGGCGCGAGCGAGAAATTAAGAAAGTTGCGGAGGCTATG 689  
Db 868 CCATCGAAGTGTGAGGCGCGAGCGAGAAATTAAGAAAGTTGCGGAGGCTATG 927  
QY 690 AGAATGATGTGGCTGCCATGAGCTCTCACTGATCCCTGGCTCAACCTGGCTCTCTAG 749  
Db 928 AGAATGATGTGGCTGCCATGAGCTCTCACTGATCCCTGGCTCAACCTGGCTCTCTAG 987  
QY 750 GTCTTTTCCAGCTTCACTCAGCGAGTCCCGCGGCTCCAGCAGGCTTACTGGGGCTG 809  
Db 988 GTCTTTTCCAGCTTCACTCAGCGAGTCCCGCGGCTCCAGCAGGCTTACTGGGGCTG 1047  
QY 810 CTCCTATAGCTCTTTTATGAGGCTCCGAGCAGGAGTGGTGCAGGTGTTATCCCG 869  
Db 1048 CTCCTATAGCTCTTTTATGAGGCTCCGAGCAGGAGTGGTGCAGGTGTTATCCCG 1107  
QY 870 CCCAGGAGTGGGCGCCATCATCGCAAGAGGGGCGAGCAATCAAAACAGCTCTCCGGT 929  
Db 1108 CCCAGGAGTGGGCGCCATCATCGCAAGAGGGGCGAGCAATCAAAACAGCTCTCCGGT 1167  
QY 930 TTGCGAGCGCTCCATCAGATTCGACCAACCGGAAACCTGACTCCAAAGTTCTGATGG 989  
Db 1168 TTGCGAGCGCTCCATCAGATTCGACCAACCGGAAACCTGACTCCAAAGTTCTGATGG 1227  
QY 990 TTATCATCACTGGACCGCGAGAGGCCCAATTCAGAGGCTCAGGGAAGATCTATGGCAAC 1049  
Db 1228 TTATCATCACTGGACCGCGAGAGGCCCAATTCAGAGGCTCAGGGAAGATCTATGGCAAC 1287  
QY 1050 TCAAGGAGGAACTTTCTTGGTCCCAAGGAGGAGTGAAGCTGGAGACCCACATAGCTG 1109  
Db 1288 TCAAGGAGGAACTTTCTTGGTCCCAAGGAGGAGTGAAGCTGGAGACCCACATAGCTG 1347  
QY 1110 TGCCAGCATCAGCAGCTGGCGGGTCAATTGGCAAGGTTGGAATAACCGTGAACGAGTTGC 1169  
Db 1348 TGCCAGCATCAGCAGCTGGCGGGTCAATTGGCAAGGTTGGAATAACCGTGAACGAGTTGC 1407

QY 1170 AGAATTTGCGGCAGCTGAGGTGGTAGTACCAGAGACAGACCCCTGATGAGAACGACC 1229  
Db 1408 AGAATTTGCGGCAGCTGAGGTGGTAGTACCAGAGACAGACCCCTGATGAGAACGACC 1467  
QY 1230 AGGTCAATCGTGAATAATCATCGGACATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCC 1289  
Db 1468 AGGTCAATCGTGAATAATCATCGGACATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCC 1527  
QY 1290 GAGACATCTGGCCCGAGGTTAAGCAGCAGCATCAGAGGAGCAGAGTAACCGGCCCAGG 1349  
Db 1528 GAGACATCTGGCCCGAGGTTAAGCAGCAGCATCAGAGGAGCAGAGTAACCGGCCCAGG 1587  
QY 1350 CACGGAGGAAGTGACACAGCCCTCTCCCTGCTTTCCTTTCAGGACAAACGCGGCAGAA 1409  
Db 1588 CACGGAGGAAGTGACACAGCCCTCTCCCTGCTTTCCTTTCAGGACAAACGCGGCAGAA 1647  
QY 1410 ATCAGAGTGTGCTCTCTCCCGCAGGCTTCCAGATGAGTGGGAATCCGGGACACNTGGGC 1469  
Db 1648 ATCAGAGTGTGCTCTCTCCCGCAGGCTTCCAGATGAGTGGGAATCCGGGACACNTGGGC 1707  
QY 1470 CGGGCTGTAGATCAGGTTTGGCCACTTGTATGAGAAAGATGTTCCAGTCAGGAACCTCGA 1529  
Db 1708 CGGGCTGTAGATCAGGTTTGGCCACTTGTATGAGAAAGATGTTCCAGTCAGGAACCTCGA 1767  
QY 1530 TCTNTCAGCCCAACACCCACCAATTCGCCCCAACACTGTNTGCCCTCGGGGTCTCAG 1589  
Db 1768 TCTNTCAGCCCAACACCCACCAATTCGCCCCAACACTGTNTGCCCTCGGGGTCTCAG 1827  
QY 1590 AAATTTAGCGCAAGGCACTTTTAACGTCGATGTTTAAAGAGCTCTCCAGGCCCCAC 1649  
Db 1828 AAATTTAGCGCAAGGCACTTTTAACGTCGATGTTTAAAGAGCTCTCCAGGCCCCAC 1887  
QY 1650 CAAGAGGTGGATCACACCTCAGTGGGAGAAATAAATAATTCCTTCAGGTTTAAAA 1708  
Db 1888 CAAGAGGTGGATCACACCTCAGTGGGAGAAATAAATAATTCCTTCAGGTTTAAAA 1946

## RESULT 3

US-10-062-674-2082/c  
; Sequence 2082, Application US/10062674  
; Publication No. US2004000559A1  
; GENERAL INFORMATION:  
; APPLICANT: Loring, Jeanne F.; Kaser, Matthew R.  
; TITLE OF INVENTION: MARKERS OF NEURONAL DIFFERENTIATION AND MORPHOGENESIS  
; FILE REFERENCE: PA-0026-1 CIP  
; CURRENT APPLICATION NUMBER: US/10/062,674  
; CURRENT FILING DATE: 2002-01-30  
; PRIOR APPLICATION NUMBER: US 09/625,102  
; PRIOR FILING DATE: 2000-07-24  
; NUMBER OF SEQ ID NOS: 2217  
; SOFTWARE: PERL Program  
; SEQ ID NO 2082  
; LENGTH: 1738  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No. US2004000559A1 457674.1  
US-10-062-674-2082

Query Match 75.2%; Score 1283.6; DB 16; Length 1738;  
Best Local Similarity 93.6%; Pred. No. 0;  
Matches 1615; Conservative 0; Mismatches 64; Indels 46; Gaps 25;  
QY 30 CCGGGGAGCCATCATGAAGCTGAATGGCCACCAAGTTGGAGAACCATGCTTGAAGGTCT 89  
Db 1735 CCAGGCAAGCCATCATGAAGCTGAATGGCCACCAAGTTGGAGAACCATGCTTGAAGGTCT 1676  
QY 90 CTTACATCCCATGAGCAGATAGCACAGGAGCTTGAATGGCGCGCCGAGGGGCTTTG 149  
Db 1675 CTTACATCCCATGAGCAGATAGCACAGGAGCTTGAATGGCGCGCCGAGGGGCTTTG 1616



150 QY GCTCTCGGGTACGCCCCGAGGGCTCACCTGTGGCAGGGGGGCCCCAGCAGC 209  
1615 Db GCTCTCGGGGTACGCCCCGAGGGCTCACCTGTGGCAGGGGGGCCCCAGCAGC 1556  
210 QY AGCAAGTGGACATCCCTTCGGGTCTCTGGTGGCCACCCAGTATGTGGGTGCCATTATG 269  
1555 Db AGCAAGTGGACATCCCTTCGGGTCTCTGGTGGCCACCCAGTATGTGGGTGCCATTATG 1496  
270 QY GCAAGGAGGGGCCACCATCCGCAACATCAAAACAGACCCAGTCCAAAGATAGACGTGC 329  
1495 Db GCAAGGAGGGGCCACCATCCGCAACATCAAAACAGACCCAGTCCAAAGATAGACGTGC 1436  
330 QY ATAGGAGGAGAACGCGAGTGCAGCTGAAAGCCATCGTGTGCATCCACCCCTGAGG 389  
1435 Db ATAGGAGGAGAACGCGAGTGCAGCTGAAAGCCATCGTGTGCATCCACCCCTGAGG 1376  
390 QY GCTGCTCTCCG-CTTGTAGATGATCTTGAGAGATTATGCATAAGAGGGCTAAGGACAC 448  
1375 Db GCTGCTCTCCGCTTGTAGATGATCTTGAGAGATTATGCATAAGAGGGCTAAGGACAC 1316  
449 QY AAAAGGCTGAGAGGTTCCCTGAAGATCCTGGCCCAATAAATCTTTGTAGGCGTCTC 508  
1315 Db AAAAGGCTGAGAGGTTCCCTGAAGATCCTGGCCCAATAAATCTTTGTAGGCGTCTC 1256  
509 QY -ATTGGCAAGGAGGACGGAACCTGAAGAA-GGTAGAGCAAGATACCGAGACAAATCA 566  
1255 Db AATTGGCAAGGAGGACGGAACCTGAAGAGGGTAGAGCAAGATACCGAGACAAATCA 1196  
567 QY CCATCTCTCGTTGCAAGACCTTACCTTTACAA-CCCTGAGAGGACATCACTGTGAA- 624  
1195 Db CCATCTCTCGTTGCAAGACCTTACCTTTACAA-CCCTGAGAGGACATCACTGTGAA 1136  
625 QY GGGGCCATCGAATCTGTCAGGG-CCGAGCAGGAATAATCAAGAAAGTTCCGGAGG 683  
1135 Db GGGGCCATCGAATCTGTCAGGGCCGAGCAGGAATAATCAAGAAAGTTCCGGAGG 1076  
684 QY CCTATGAGATGATGTGCTGCCATGAGC-----TCTACCTGATCCCTGGGCTGAAC 737  
1075 Db CCTATGAGATGATGTGCTGCCATGAGCCTGCGAGTCTCACCTGATCCCTGGGCTGAAC 1016  
738 QY TGGCTGCTGTAGGTCTTTCCAGGTCATCCAGGCGAGTCCCGCGCTCCAGCAGG 797  
1015 Db TGGCTGCTGTAGGTCTTTCCAGGTCATCCAGGCGAGTCCCGCGCTCCAGCAGG 956  
798 QY TTACTGGGGCTGCTCCCTATAGCTCTTTATGAGGCTCCGAGCAGAGATGTGAGG 857  
955 Db TTACTGGGGCTGCTCCCTATAGCTCTTTATGAGGCTCCGAGCAGAGATGTGAGG 896  
858 QY TGTTTATCCCGCCAGCAGTGGCGGCTCATCGGCAAGAGGGGAGCAGATCAAC 917  
895 Db TGTTTATCCCGCCAGCAGTGGCGGCTCATCGGCAAGAGGGGAGCAGATCAAC 836  
918 QY AGCT--CTCCGGTTTGCGAGCGCTCC-ATCAAGATTGCACCCAGCAACACCTGACT 974  
835 Db CAGCTCTCCCGTTTGCGAGCGCTCCAAATCAAGATTGCACCCAGCAACACCTGACT 775  
975 QY CAAAAGTTTGTATGTTATATCATCTGAGCCGAGG-CCCAATTCAGGCTCAGGGA 1033  
775 Db CAAAAGTTTGTATGTTATATCATCTGAGCCAGGAGGCCCCCAATTCAGGCTCAGGGA 716  
1034 QY AGAATCTATGG-CAAACTCAGGAGGAACTTTT-TGGTCCCAAGAGGAGTGAAGC 1091  
715 Db AGAATCTATGGCAAACTCAGGAGGAACTTTCTTGGGTTCCCAAGAGGAGTGAAGC 656  
1092 QY TGGAGACCCACATACGTGTGCAGCATCAGCAGTGGCCGG--TCATTGGCAAGGTGG 1149  
655 Db TGGAGACCCACATACGTGTGCAGCATCAGCAGTGGCCGGCTCANTTGGCCAAAGGTGG 596  
1150 QY AAAAAC-GGTGACAGTTGAGATTTGAGGAGGCTGAGTGGTAG----TACCAGA 1204  
595 Db AAAAACGGGTGACAGTTGAGATTTGAGGAGGCTGAGTGGTAGTACCAACAGA 536  
1205 QY GACCAGACCCCTGATGAGAACGACGAGGTATCGTGAATAATCATCGGAC-ATTCTATG 1263

535 Db CCAGACCCCTGTATGAAGAACGACAGGTCTATCGTGAATAATCATCGGACAAATTTCTATGC 476  
1264 QY CAGTCAGATGGCTCAACGGAA-GATCCGAGACATCTCTGGCCAGGTTAAGCAGCAGCATC 1322  
475 Db CAGTCAGATGGCTCAACGGAAAGGATCCGAGACATCTCTGGCCAGGTTAAGCAGCAGCATC 416  
1323 QY AGAAGGGACAGATTAACGAGCCAGGACGAGGAGTGAACGAGGCTC-CTGTGCTCC 1381  
415 Db AGAAGGGACAGATTAACGAGCCAGGACGAGGAGTGAACGAGGCTCCTGCTGTCTCC 356  
1382 QY TTNGAGTCCAGGACAAACAGGGCAGAAATCGAGTGTGTCTCTCCCGGAGGCTGAG 1441  
355 Db TTCGAGTCCAGGACAAACAGGGCAGAAATCGAGTGTGTCTCTCCCGGAGGCTGAG 296  
1442 QY AATGAGTGGAAATCCGGACAC-NTGGCCGGGCTGTGTAGATCAGGTTTGGCCAC-----T 1495  
295 Db AATGAGTGGAAATCCGGACACACTGGCCGGGCTGTGTAGATCAGGTTTGGCCACACTGAT 236  
1496 QY TCAATTGAGAAAGATGTTCCAGTGAGGAACCTGTATCTTNTCAGCCCC-----AAACACCC 1549  
235 Db TGAGAAACGATGTTCCAGTGAAGAAACCTGTATCTCTTAGCCCCCACTACACCCAC 176  
1550 QY ACCCAATTGG-----CCCAACACTGTTTGGCCCTCGGGGTGT-CAGAAATTTAGCGCAAG- 1604  
175 Db ACCCAATTGGCCCAACAACTGTCTGGCCCTCGGGGTGTCCAGAAATCTAGCGCAAGA 116  
1605 QY GCACTTTTAAAGCT-GGATTTGTTTAAAGAGCTCTCCAGGCCCCCAAGAGGTTGGATC 1663  
115 Db GCACTTTTAAAGCTGGGATTTGTTTAAAGAGCTCTCCAGGCCCCCAAGAGGTTGGATC 56  
1664 QY ACACCTCAGTGGGAGAGAAATAAATTTCCCTTCAGGTTTTTAAAA 1708  
55 Db ACACCTCAGTGGGAGAGAAATAAATTTCCCTTCAGGTTTTTAAAA 11

RESULT 4  
US-09-873-637-1  
; Sequence 1, Application US/09873637  
; Patent No. US20020061543A1  
; GENERAL INFORMATION:  
; APPLICANT: Rossi, Jeffrey  
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN  
; FILE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE  
; FILE REFERENCE: 960296.95131  
; CURRENT APPLICATION NUMBER: US/09/873,637  
; CURRENT FILING DATE: 2001-06-04  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 2224  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-873-637-1

Query Match 58.2%; Score 1165; DB 9; Length 2224;  
Best Local Similarity 82.8%; Pred. No. 0;  
Matches 1411; Conservative 0; Mismatches 260; Indels 33; Gaps 6;

30 CCGGGGAGCCATCATGAAGCTGAATGGCCACACAGTGTGGAGAACCATCCCTGAAGGTCT 89  
525 CCAGGCAAGCTATCATGAAGCTAAATGGCCATCACTGGAGAACCATCCCTGAAGGTCT 584  
90 CTTACATCCCATGATGAGCAGATAGCAGAGGACCTGAGAAATGGGCGGAGGGCTTTG 149  
585 CTTACATCCCTCATGAGCAGATTAACAGAGTCTCTGAGAAATGGGCGTCTGTGAGGCTTTG 644  
150 GCTCTCGGGGTGAGCCCGCCAGGGGTCACTGTGGCAGCGGGGGCCCCAGCAAGCAGC 209  
645 GGTCTCGGGGCCAGCCCCGGCAGGGTGGCCCGTGGCAGCAGGGGCTCCAGCAAGCAGC 704  
210 AGCAAGTGAACATCCCTTTCGGCTCTGTGGTCCGCCACCGAGTATGTGGTGCCATTATG 269



705 AGCCAGTGGACATCCCTCTCCGGCTCCTGGTCCCTAGCCAGTATGTAGGCGCTATCATTTG 764  
270 GCAGAGGGGGCCACCATCCGCAACATCATCACAACACAGACCCAGTCCCAAGTATAGAGTGC 329  
765 GCAGAGGGGTGCCACCATCCGAACATCATCAAAACACAGACGAGTCCAAATATAGACGTGC 824  
330 ATAGGAAGGAGAACGAGGTGCAGCTGAAAGAGCCATCAGTGTGCACTCCACCCCTGAGG 389  
825 ATAGGAAGGAGAAATCGGGGCTCGGAGAGGCCATCAGCGTGCATTCAACCCCTGAAG 884  
390 GCTCTCTCCGCTGTGAAGTATGATCTTGAAGTATGATGATTAAGAGAGGCTAAGGACACCA 449  
885 GCTCTCTCCGCTGTGAAGTATGATCTTGAAGTATGATGATTAAGAGAGGCTAAGGACACCA 944  
450 AAAAGGCTGACAGGTTCCCTCCCTGAAGATCTTGGGCCCATATAACTTTGTAGGGGCTCTCA 509  
945 AAAAGGAGATGAAGTTCCCTCCCTGAAGATCTTGGGCTCATTAACAATTTGTGGGCACTCA 1004  
510 TTGGCAGGAAGAGCGGAACCTGAAGAGGTGAGCAAGATACCGAGACAAAATACCA 569  
1005 TTGGCAGGAAGAGCGGAACCTGAAGAGGTGAGCAAGATACCGAGACAAAATACCA 1064  
570 TCTCTCTGTTCAGAGCTTACCTTACACCTGAGAGACCATCACTGTGAAGGGG 629  
1065 TCTCATGCTCCAGACCTCAGCTCTATTAACCTGAGAGACCATCACTGTGAAGGGG 1124  
630 CCATCGAGAATTTGTCAGGCGCCGAGCAGGAAATATTAAGAAAGTTTCGGAGGCGCTATG 689  
1125 CCATTGAGAACTGTTCAGGCGCCGAGCAGGAGATCATGAAGAAAGTTTCGAGAGGCTTAAG 1184  
590 AGAATGATGGCTGCCATGAGC-----TCTACCTGATCCCTGGCTGAACTGGCTG 743  
1185 AGAAGCAGTGGCGCCATGAGTGTGCAAGTCCCACTCATCCCTGGGCTTAACCTGGCTG 1244  
744 CTGTAGGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCCCTCCAGCAGCGCTTACTG 803  
1245 CTGTAGGTCTTTTCCAGCTTCATCCAGCGCTGTCCTCTCCAGCAGTGTCACTG 1304  
804 GGGCTGCTCCCTATAGCTCTTTATGAGGCTCCCGAGCAGGAGATGTCAGGTGTTTA 863  
1305 GGGCTGCTCCCTATAGCTCTTTATGAGGCTCCCGAGCAGGAGATGTCAGGTGTTTA 1364  
864 TCCCGCCAGCGAGTGGGCGCCATCATCGCAAGAGGGGCGACACATCAAGCAGCTCT 923  
1365 TCCCGCCAGCGTGTGGCGCCATCATGCAAGAGGGGCGACACATCAAGCAGCTCT 1424  
924 CCGGTTTTCAGCGCTTCATCAAGATTCACCAAGAGTTCACCAAGAGTTCACCAAGTTC 983  
1425 CCGGTTTTCAGCGCTTCATCAAGATTCACCAAGAGTTCACCAAGAGTTCACCAAGTTC 1484  
984 GTATGTTTATCATCACTGAGCGCGCAGAGGCGCAATTCAGGCTCAGGAGATCTATG 1043  
1485 GAATGGTCTCATCACTGAGCGCGCGAGGCTCAGTTCAGGCTCAGGAGAAATTTATG 1544  
1044 GCAAACTCAAGAGAGAACTTCTTTGGTCCCAAGAGGAGTGAAGCTGGAGACCCACA 1103  
1545 GCAAACTCAAGAGAGAACTTCTTTGGTCCCAAGAGGAGTGAAGCTAGAGACCCACA 1604  
1104 TACGTGTCCAGCATCAGCAGCTGCGCGGCTCATTTGGCAAGGTGGAAACCGTGAAG 1163  
1605 TACGGGTTCCGGCTTCAGCAGCGCGCGCTCATCGGCAAGGGCGCAAAACCGTGAATG 1664  
1164 AGTTGCAGAAATTTGAGCGCAGCTGAGTGGTGTAGTACCAAGAGACAGACCCCTCATGAGA 1223  
1665 AGTGCAGAACTTGAATCAGCTGAGTGGTGTAGTGGCCAGAGACAGACCCCGATGAGA 1724  
1224 ACGACCAAGTCAATCGTGAATATCATCGGACATTTCTATGCGAGTCAGATGGCTCAACGGA 1283  
1725 ACGACCAAGTCAATGTTAAGATCATCGGACATTTCTATGCGAGCGAGATGGCTCAGCGGA 1784  
1284 AGATCCGAGAGATCTGCGCCAGGTAGCAGCGCATTCAGAGGAGCAGAGTACCCAGG 1343  
1785 AGATCCGAGAGATCTGCGCTCAAGTTAAGCAACAGCACAGAGGGACAGAGAACCTGG 1844

1344 CCCAGCAGGAGAGTGA-CCAGCCCTCCTCTGCTCCCTTNGAGTCCAGGACAAACG 1402  
1845 CCCAGCAGGAGAGTGA-CCCGCCCTCCTCTGCTCCCTTNGAGTCCAGGACAAACG 1904  
1403 GGAGAA-----ATCAGAGTGTGCTCTCCCGGCGAGGCTGAG 1441  
1905 AGGAACACAGAACTGGAGGGGGGTGGAGGGCGGTGTGTTTTTCCAGCAGGCGTGA 1964  
1442 AATGAGTGGAAATCCGGGACACNTGGCGGGCTGTAGATCAGGTTTCCCTCTGATG 1501  
1965 AATGAGTGGAAATCAGGG-CATTTGGGCTGCTGAGATCAGGTTTGCACACTGTATTG 2023  
1502 AGAAGATGTTCCAGTGAAGAACCTGTATCTNTCAGCCGCCAACACCCCAATTTGGCC 1561  
2024 AGAACAATGTTCCAGTGAAGAACTCTGATCTCTCGCCCCCAA--TTGAGCCAGCTGGCCA 2081  
1562 CAACACTGNTGCCCTCGGGGTGTGAGAAATNTAGCAAGGACACTTTTAAACGTGA 1621  
2082 CAGCCACACCTTGGAAATATCACCATTGCAATCATAGCTTGGTGTCTTTTAAACGTGA 2141  
1622 TTGTTTAAAGAACTCTCAGGCGCCACCAAGAGGGTGGATCACACCTCAGTGGGAAGAA 1681  
2142 TTGTCT--TGAAGTCTCCAGCTCCATGGAAGATGGGTCCAGATCCAGTGGGAAGAG 2199  
1682 AATAAAATTTCTTCAGGTTTTTA 1705  
2200 AATAAAATTTCTTCAGGTTTTTA 2223

RESULT 5

US-09-764-864-327  
; Sequence 327, Application US/09764864  
; Patent No. US20020132753A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: P223  
; CURRENT APPLICATION NUMBER: US/09/764,864  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1792  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 327  
; LENGTH: 1182  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-864-327

Query Match 63.5%; Score 1085.4; DB 9; Length 1182;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 1102; Conservative 0; Mismatches 6; Indels 6; Gaps 1;

Qy 601 CCTGAGAGGACCATCACTGTGAAGGGGCCCATCGAGAATTTGTCAGGCGCCGAGCAGGA 660  
Db 18 CGCTGAGAGGACCATCACTGTGAAGGGGCCCATCGAGAATTTGTCAGGCGCCGAGCAGGA 77  
Qy 661 AATAATGAAGAAAGTTCCGGAGCCCTATGAGATGATGTGGCTGCCATGAGC-----TC 714  
Db 78 AATAATGAAGAAAGTTCCGGAGCCCTATGAGATGATGTGGCTGCCATGAGCCTGAGTC 137  
Qy 715 TCACCTGATCCCTGGGCTGAACTGGCTGCTGTAGTGTCTTTTCCAGCTTCATCCAGCGC 774  
Db 138 TCACCTGATCCCTGGGCTGAACTGGCTGCTGTAGTGTCTTTTCCAGCTTCATCCAGCGC 197  
Qy 775 AGTCCCGCGCTCCCGAGCAGGTTACTGGGGTGTCTCCCTATAGCTCCCTTTATGACAGG 834  
Db 198 AGTCCCGCGCTCCCGAGCAGGTTACTGGGGTGTCTCCCTATAGCTCCCTTTATGACAGG 257  
Qy 835 TCCCGAGCAGGAGATGGTGCAGGTGTTATCCCGGCCAGGAGTGGGGCCATCATCGG 894  
Db 258 TCCCGAGCAGGAGATGGTGCAGGTGTTATCCCGGCCAGGAGTGGGGCCATCATCGG 317

```
QY 895 CAAGAAGGGGAGCAGCATCAACAGCTCTCCGGTTTCCAGCGCCTCCATCAGATTGC 954
Db 318 CAAGAAGGGGAGCAGCATCAACAGCTCTCCGGTTTCCAGCGCCTCCATCAGATTGC 377
QY 955 ACCACCCGAAACACCTGACTCAAAAGTTTCGATGTTATCATCTACTGACCGCAGAGGC 1014
Db 378 ACCACCCGAAACACCTGACTCAAAAGTTTCGATGTTATCATCTACTGACCGCAGAGGC 437
QY 1015 CCAATTCAAGCTCAGGGAAGAACTTATGGCAAACTCAAGGAGAGAACTTCTTTGTTCC 1074
Db 438 CCAATTCAAGCTCAGGGAAGAACTTATGGCAAACTCAAGGAGAGAACTTCTTTGTTCC 497
QY 1075 CAAAGAGGAAAGTGAAGCTGGAGACCCATACGTGTGCGACATCAGCAGCTGCCCGGT 1134
Db 498 CAAAGAGGAAAGTGAAGCTGGAGACCCATACGTGTGCGACATCAGCAGCTGCCCGGT 557
QY 1135 CATTGGCAAAAGTGGAAAAACGGTGAACGAGTTGAGAAATTTGACGGCAGCTGAGGTGGT 1194
Db 558 CATTGGCAAAAGTGGAAAAACGGTGAACGAGTTGAGAAATTTGACGGCAGCTGAGGTGGT 617
QY 1195 AGTACCAAGACAGACAGCCCTGTATGAGAACGACAGCTCATCTGTGAAATCATCGGACA 1254
Db 618 AGTACCAAGACAGACAGCCCTGTATGAGAACGACAGCTCATCTGTGAAATCATCGGACA 677
QY 1255 TTTCCTATCCAGTCAAGTGGTCAACGGAAGATCCGAGACATCTCTGCCAGGTTAAAGCA 1314
Db 678 TTTCCTATCCAGTCAAGTGGTCAACGGAAGATCCGAGACATCTCTGCCAGGTTAAAGCA 737
QY 1315 GCAGCATCAGAAAGGACAGTAAACCAGGCCAGCGAGGAAGTACCGCCCTCC 1374
Db 738 GCAGCATCAGAAAGGACAGTAAACCAGGCCAGCGAGGAAGTACCGCCCTCC 797
QY 1375 CTGTCCCTTNGAGTCAGGACAAACAAACGGGAGAGAAATCGAGAGTGTCTCTCCCGGCAG 1434
Db 798 CTGTCCCTTNGAGTCAGGACAAACAAACGGGAGAGAAATCGAGAGTGTCTCTCCCGGCAG 857
QY 1435 GCCTGAGAAATGATGGGAAATCGGGACACNTGGCGCGGCTGTAGATCAGTTTGGCCAC 1494
Db 858 GCCTGAGAAATGATGGGAAATCGGGACACNTGGCGCGGCTGTAGATCAGTTTGGCCAC 917
QY 1495 TTGATTGAGAAAGATGTTCCAGTGAAGAAACCTGATCTNTCAGCCCAACCAACCCCA 1554
Db 918 TTGATTGAGAAAGATGTTCCAGTGAAGAAACCTGATCTNTCAGCCCAACCAACCCCA 977
QY 1555 ATTGGCCCAACACTGNTGCCCTCGGGGTGTCAAGAAATNTAGCGAAGGCACTTTTAA 1614
Db 978 ATTGGCCCAACACTGNTGCCCTCGGGGTGTCAAGAAATNTAGCGAAGGCACTTTTAA 1037
QY 1615 ACGTGGATGTTTAAAGAGCTCTCCAGGCCCAACCAAGAGGTTGATCACTCAGTG 1674
Db 1038 ACGTGGATGTTTAAAGAGCTCTCCAGGCCCAACCAAGAGGTTGATCACTCAGTG 1097
```

## RESULT 6

```
US-09-764-864-745
; Sequence 745, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P7223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 745
; LENGTH: 1129
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
US-09-764-864-745
Query Match 62.6%; Score 1069.8; DB 9; Length 1129;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1084; Conservative 3; Mismatches 5; Indels 6; Gaps 1;
QY 617 ACTGTCAAGGGGGCCATCGAGAAATTTGTCAGGGCCGAGCAGGAATAATGAGAAAGCTT 676
Db 1 ACTGTCAAGGGGGCCATCGAGAAATTTGTCAGGGCCGAGCAGGAATAATGAGAAAGCTT 60
QY 677 CCGGAGGCCCTATGAGAATGATGTGGCTGCCATGAGC-----TCTACCTGATCCCTGGC 730
Db 61 CCGGAGGCCCTATGAGAATGATGTGGCTGCCATGAGCCTGCAGTCTCACCTGATCCCTGGC 120
QY 731 CTGAACCTGGCTGCTGTAGTCTCTTTCCAGGTTTATCCAGGGAGTTCGCGCGCTCCC 790
Db 121 CTGAACCTGGCTGCTGTAGTCTCTTTCCAGGTTTATCCAGGGAGTTCGCGCGCTCCC 180
QY 791 AGCAGCGTTTACTGGGGCTGCTCCCTATAGTCTCTTTATGCAAGGCTCCGAGCAGAGATG 850
Db 181 AGCAGCGTTTACTGGGGCTGCTCCCTATAGTCTCTTTATGCAAGGCTCCGAGCAGAGATG 240
QY 851 GTGCAAGGTTTATCCCGGCCAGCAGTGGGGCCATCATCGGCAAGAGGGGAGCAG 910
Db 241 GTGCAAGGTTTATCCCGGCCAGCAGTGGGGCCATCATCGGCAAGAGGGGAGCAGC 300
QY 911 ATCAACAGCTCTCCCGGTTTCCAGCGCTCCATCAAGATTGCACACCCGAAACACCT 970
Db 301 ATCAACAGCTCTCCCGGTTTCCAGCGCTCCATCAAGATTGCACACCCGAAACACCT 360
QY 971 GACTCCAAAGTTTGTATGTTTATCATCTACTGACCGCAGAGGGCCAAATTCAGGCTCAG 1030
Db 361 GACTCCAAAGTTTGTATGTTTATCATCTACTGACCGCAGAGGGCCAAATTCAGGCTCAG 420
QY 1031 GGAAGAACTATGCGAAACTCAAGAGAGAACTTCTTTGGTCCCAAGGAGAGTGAAG 1090
Db 421 GGAAGAACTATGCGAAACTCAAGAGAGAACTTCTTTGGTCCCAAGGAGAGTGAAG 480
QY 1091 CTGGAGAGCCACATACGTGTGCGACATCAGCAGCTGCGCGGTTCATTGGCAAGAGTGA 1150
Db 481 CTGGAGAGCCACATACGTGTGCGACATCAGCAGCTGCGCGGTTCATTGGCAAGAGTGA 540
QY 1151 AAAACGGTGAACGAGTTGAGAAATTTGAGCGCAGCTGAGGTGTTAGTACCAAGAGCAG 1210
Db 541 AAAACGGTGAACGAGTTGAGAAATTTGAGCGCAGCTGAGGTGTTAGTACCAAGAGCAG 600
QY 1211 ACCCTGTATGAGAACGACACAGGTCTATCGTGAATAATCATCGGACATTTCTATGCCAGT 1270
Db 601 ACCCTGTATGAGAACGACACAGGTCTATCGTGAATAATCATCGGACATTTCTATGCCAGT 660
QY 1271 ATGGTCTCAACGGAAGATCCGAGACATCTGCGCCAGGTTAAGCAGCAGCATCAGAGGGA 1330
Db 661 ATGGTCTCAACGGAAGATCCGAGACATCTGCGCCAGGTTAAGCAGCAGCATCAGAGGGA 720
QY 1331 CAGAGTAAACGAGGCCAGCGAGGAGGAGTACAGCGCCCTCCCTGTCCTTNGAGTCC 1390
Db 721 CAGAGTAAACGAGGCCAGCGAGGAGGAGTACAGCGCCCTCCCTGTCCTTNGAGTCC 780
QY 1391 AGGACAAACACGGGAGAAATCGAGAGTGTCTCTCCCGCGCAGGCTTGAGAAATGAGTGG 1450
Db 781 AGGACAAACACGGGAGAAATCGAGAGTGTCTCTCTCCCGCGCAGGCTTGAGAAATGAGTGG 840
QY 1451 GAATCCGGGACACNTGGCGCGGCTGTAGATCAGGTTTGCCTTCTGATTCAGAAAGATG 1510
Db 841 GAATCCGGGACACNTGGCGCGGCTGTAGATCAGGTTTGCCTTCTGATTCAGAAAGATG 900
QY 1511 TTCCAGTGAAGAACCTCTGATCTNTCAGCGCCCAACACACCCCAATTTGGCCCAACATGT 1570
Db 901 TTCCAGTGAAGAACCTCTGATCTCTCAGCGCCCAACACACCCCAATTTGGCCCAACATGT 960
QY 1571 NTGCCCCCTCGGGGTGTGAGAAATTTAGCGCAGGCACTTTTAAACGTGGATTGTTTAA 1630
```

|    |      |  |      |
|----|------|--|------|
| Db | 961  | CTGCCCTCGGGTGTCTAGCCACAGGCATTTTAAACGTGGATGTTTAA            | 1020 |
| Qy | 1631 | GAAGCTCTTCAGGCCCCACCAAGAGGGTGGATCACACCTCAGTGGGAAGAAAATTAAT | 1690 |
| Db | 1021 | GAAGCTCTTCAGGCCCCACCAAGAGGGTGGATCACACCTCAGTGGGAAGAAAATTAAT | 1080 |
| Qy | 1691 | TTCCCTTCAGGTTTTAAAA  | 1708 |
| Db | 1081 | TTCCCTTCAGGTTTTAAAA  | 1098 |

## RESULT 7

```

US-09-735-705-347
; Sequence 347, Application US/09735705
; Patent No. US20020052329A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Ligu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy R.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS
; TITLE OF INVENTION: AND DIAGNOSIS OF LONG
; FILE REFERENCE: 210121.455C14
; CURRENT APPLICATION NUMBER: US/09/735,705
; CURRENT FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 419
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 347
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-735-705-347

```

|                           |       |                     |           |              |
|---------------------------|-------|---------------------|-----------|--------------|
| Query Match               | 34.1% | Score 582.2;        | DB 9;     | Length 1740; |
| Best Local Similarity     | 66.2% | Pred. No. 6.9e-171; |           |              |
| Matches 858; Conservative | 0;    | Mismatches 433;     | Indels 6; | Gaps 1;      |

|    |     |  |     |
|----|-----|--|-----|
| Qy | 37  | AGCCATCATGAGCTGAAATGCGCCACCAAGTTGGAGAACCATGCCCTCGAAGGCTCTCCTACAT | 96  |
| Db | 402 | AGCACTAGACAAACTGAATGGATTTTCAGTTTAGAGAAATTTCACCTTGAAAGTAGGCTTAT   | 461 |
| Qy | 97  | CCCCGATGAGCAGATAGCACAGGGAACCTTGAGAAATGGGCGCGAGGGGGCTTTTGGCTCTCG  | 156 |
| Db | 462 | CCCTGATGAACGGCGGCCCAAGCAAAACCCCTTGCAGCAGCCCGAGGTGCCCGGGGCT       | 521 |
| Qy | 157 | GGGTGAGCCCGCCAGGGCTCACCTGTGGCAGCGGGGCCCCAGCCAGCAGCAAGT           | 216 |
| Db | 522 | TGGCAGAGGGGCTCCTCAAGGCAGGGGCTCTCCAGGATCCGTATCCAAGCAGAAACCATG     | 581 |
| Qy | 217 | GGACATCCCCCTTCGGCTCTGGTGTCGCCACCCAGTATGTGGTGCCATTATTGGCAGGA      | 276 |
| Db | 582 | TGATTTGGCTCTGGCGCTGCTGGTTGCCACCCCAATTTGTTGGAGCCATCATAGGAAAAGA    | 641 |
| Qy | 277 | GGGGGCCACCATCCGCAACATCAAAAACAGACCCAGTCCCAAGATAGACGTGCATAGAA      | 336 |
| Db | 642 | AGGTGCCACCATTCGGAACATCAACAAACAGACCCAGTCTAAATTCGATGTCCACCGTAA     | 701 |
| Qy | 337 | GGAGAACACAGTGCAGCTGAAAAGCCATCAGTGTGCATCTCCACCCCTGAGGGGTGTC       | 396 |
| Db | 702 | AGAAATCGGGGGCTGTGAGAGTCAATTACTATCTCTCTACTCTCTGAGGCACCTC          | 761 |
| Qy | 397 | CTCCGCTTGTAAAGTATCTTGGAGATTATGCATAAAGAGGCTAAGGCACACAAAACGGC      | 456 |
| Db | 762 | TGGGCTTGTAAGTCTATCTTGGAGATTATGCATAAAGAAAGCTCAAGATATAAATTCAC      | 821 |

```
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 347
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-850-716A-347

Query Match      34.1%; Score 582.2; DB 9; Length 1740;
Best Local Similarity 66.2%; Pred. No. 6.9e-171;
Matches 858; Conservative 0; Mismatches 433; Indels 6; Gaps 1;

QY 37 AGCCATCATGAAGCTGAATGGCCACCACTGAGTGGAGAACCATGCCCTGAGGCTCTCTACAT 96
Db 402 AGCACTAGACAAACTGAATGGATTTCAGTTAGAGAAATTTCACTTGAAGTAGGCTATAT 461
QY 97 CCCCAGATGAGCAGATAGCACAGGACCTTGAGAAATGGGCGCCGAGGGGCTTTGGCTCTCG 156
Db 462 CCGTGATGAACGGCGCGCCAGCAAAACCCCTTCAGCAGCCCCGAGGTCGCCGGGGCT 521
QY 157 GGGTCAGCCCCCGCAGGGCTCACCCTGTGGCAGCGGGGGCCCCAGCCAGCAGCAGAGT 216
Db 522 TGGGCAGAGGGGCTCCTCAAGCAGGGGTCTCCAGGATCCGATATCCAAAGCAGAAACCATG 581
QY 217 GGACATCCCTTCGGCTCCTGGTGCCCAACCCAGTATGTGGTGCCATTATTGGCAAGGA 276
Db 582 TGATTGGCTCTGGCTCTGGTGTCCCAACCCAAATTTGTTGAGCCATCATAGAAAGA 641
QY 277 GGGGCGCCACCATCCGCAACATCAAAAACAGCCAGTCCAGATAGACGTGCAATGAA 336
Db 642 AGGTGCCACCATTCGGAACATCACAAACAGACCCAGTCTAAATCGATGTCACCGTAA 701
QY 337 GGAGAACCGAGTGGAGCTGAAAGCCATCAGTGTGCACCTCCACCCCTGAGGCTCTC 396
Db 702 AGAAATCGGGGGCTGCTGAGAGTGAATTAATCTCTACTCTGAGGACCTC 761
QY 397 CTCGCTTGTGAATGATCTTTGGAGATATGATCAATAAGAGGCTAAGGACACCAAAACGGC 456
Db 762 TCGGCTTGTGAATGATCTTTGGAGATATGATCAATAAGAGGCTAAGGACACCAAAATTCAC 821
QY 457 TGACAGGTTCCCTGAGATCCTGGCCCAATAAATCTTGTAGGGGCTCTCATTTGCCAA 516
Db 822 AGAAGAGATCCCTTGAAGATTTAGCTCAATAAATCTTGTGGAGCTCTTATGGTAA 881
QY 517 GGAAGGACGGAACCTGGAAGAGGTAGACAGATACCGAGACAAAAATCACCATCTCTC 576
Db 882 AGAAGGAAGAATCTTAAATAAATTTAGCAAGACACAGACATTAATCAGATATCTCC 941
QY 577 GTTCGAGACCTTACCCTTTACACCTTGAGAGACCATCAGTGTGAAGGGGCGCATCGA 636
Db 942 ATTGAGGAATGACGCTGTATATCCAGAACGCACTATTACAGTTAAAGGCAATGTGA 1001
QY 637 GAATTTGTGACGGCGCAGGAGAAATAATGAAGAAATTCGGAGGCTTATGAGAAATGA 696
Db 1002 GACATGTGCCAAAGCTGAGGAGGAGATCATGAAGAAATCAGGAGTCTTATGAATGA 1061
QY 697 TGTGGCTGCCATGA-----GCTCTACCTGATCCCTGGCTGAACTGCTGCTGAGG 750
Db 1062 TATTGCTTCTATGAATCTTCAAGCACAATTAATTCCTGGATTAATCTGAACGCTTGGG 1121
QY 751 TCTTTTCCAGCTTCATCCAGCGAGTCCGCGGCTCCAGCAGGCTTACTGGGCTGC 810
Db 1122 TCTGTTCCACCCACTTCAGGATGCGCACTCCACCTCAGGCGCCCTTCAGCCATGAC 1181
QY 811 TCCCTATAGTCTCTTTATGAGGCTCCCGAGCAGGAGATGCTGAGGTGTTTATCCCGC 870
Db 1182 TCCCTCCCTACCGCAGTTTGAAGCAATCAGAAACGAGACTGTTTCACTGTTTATCCAGC 1241
QY 871 CCAGGCAGTGGGCGCCATCATCGCAGAGAGGGGCGCAGCATCAAAAGCTCTCCGGTT 930
Db 1242 TCTATCAGTGGTGCCATCATCGCAGAGAGGGGCGCAGCATCAAGAGCTTTCTCGCTT 1301

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C16
; CURRENT APPLICATION NUMBER: US/09/897,778
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 347
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-897-778-347

Query Match      34.1%; Score 582.2; DB 9; Length 1740;
Best Local Similarity 66.2%; Pred. No. 6.9e-171;
Matches 858; Conservative 0; Mismatches 433; Indels 6; Gaps 1;

QY 37 AGCCATCATGAAGCTGAATGGCCACCACTGAGTGGAGAACCATGCCCTGAGGCTCTCTACAT 96
Db 402 AGCACTAGACAAACTGAATGGATTTCAGTTAGAGAAATTTCACTTGAAGTAGGCTATAT 461
QY 97 CCCCAGATGAGCAGATAGCACAGGACCTTGAGAAATGGGCGCCGAGGGGCTTTGGCTCTCG 156
Db 462 CCGTGATGAACGGCGCGCCAGCAAAACCCCTTCAGCAGCCCCGAGGTCGCCGGGGCT 521
QY 157 GGGTCAGCCCCCGCAGGGCTCACCCTGTGGCAGCGGGGGCCCCAGCCAGCAGCAGAGT 216
Db 522 TGGGCAGAGGGGCTCCTCAAGCAGGGGTCTCCAGGATCCGATATCCAAAGCAGAAACCATG 581
QY 217 GGACATCCCTTCGGCTCCTGGTGCCCAACCCAGTATGTGGTGCCATTATTGGCAAGGA 276
Db 582 TGATTGGCTCTGGCTCTGGTGTCCCAACCCAAATTTGTTGAGCCATCATAGAAAGA 641
QY 277 GGGGCGCCACCATCCGCAACATCAAAAACAGCCAGTCCAGATAGACGTGCAATGAA 336
Db 642 AGGTGCCACCATTCGGAACATCACAAACAGACCCAGTCTAAATCGATGTCACCGTAA 701
QY 337 GGAGAACCGAGTGGAGCTGAAAGCCATCAGTGTGCACCTCCACCCCTGAGGCTCTC 396
Db 702 AGAAATCGGGGGCTGCTGAGAGTGAATTAATCTCTACTCTGAGGACCTC 761
QY 397 CTCGCTTGTGAATGATCTTTGGAGATATGATCAATAAGAGGCTAAGGACACCAAAACGGC 456
Db 762 TCGGCTTGTGAATGATCTTTGGAGATATGATCAATAAGAGGCTAAGGACACCAAAATTCAC 821
QY 457 TGACAGGTTCCCTGAGATCCTGGCCCAATAAATCTTGTAGGGGCTCTCATTTGCCAA 516
Db 822 AGAAGAGATCCCTTGAAGATTTAGCTCAATAAATCTTGTGGAGCTCTTATGGTAA 881
QY 517 GGAAGGACGGAACCTGGAAGAGGTAGACAGATACCGAGACAAAAATCACCATCTCTC 576
Db 882 AGAAGGAAGAATCTTAAATAAATTTAGCAAGACACAGACATTAATCAGATATCTCC 941
QY 577 GTTCGAGACCTTACCCTTTACACCTTGAGAGACCATCAGTGTGAAGGGGCGCATCGA 636
Db 942 ATTGAGGAATGACGCTGTATATCCAGAACGCACTATTACAGTTAAAGGCAATGTGA 1001
QY 637 GAATTTGTGACGGCGCAGGAGAAATAATGAAGAAATTCGGAGGCTTATGAGAAATGA 696
Db 1002 GACATGTGCCAAAGCTGAGGAGGAGATCATGAAGAAATCAGGAGTCTTATGAATGA 1061
QY 697 TGTGGCTGCCATGA-----GCTCTACCTGATCCCTGGCTGAACTGCTGCTGAGG 750
Db 1062 TATTGCTTCTATGAATCTTCAAGCACAATTAATTCCTGGATTAATCTGAACGCTTGGG 1121
QY 751 TCTTTTCCAGCTTCATCCAGCGAGTCCGCGGCTCCAGCAGGCTTACTGGGCTGC 810
Db 1122 TCTGTTCCACCCACTTCAGGATGCGCACTCCACCTCAGGCGCCCTTCAGCCATGAC 1181
QY 811 TCCCTATAGTCTCTTTATGAGGCTCCCGAGCAGGAGATGCTGAGGTGTTTATCCCGC 870
Db 1182 TCCCTCCCTACCGCAGTTTGAAGCAATCAGAAACGAGACTGTTTCACTGTTTATCCAGC 1241
QY 871 CCAGGCAGTGGGCGCCATCATCGCAGAGAGGGGCGCAGCATCAAAAGCTCTCCGGTT 930
Db 1242 TCTATCAGTGGTGCCATCATCGCAGAGAGGGGCGCAGCATCAAGAGCTTTCTCGCTT 1301

RESULT 9
US-09-897-778-347
; Sequence 347, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C16
; CURRENT FILING DATE: 2001-08-28
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 347
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-897-778-347
```

Db 582 TGATTTGCTCTGCGCTGCTGGTTCCACCACCCATTTGTTGGAGCCATCATAGGAAGA 641  
Qy 277 GGGGGCCACCATCCGCGACATCAACAAACAGACCCAGTCCAAAGATAGAGTGTGATAGGAA 336  
Db 642 AGGTGCCACCATTCGGAACATCACCAACAGACCCAGTCTAAATTCGATGTCCACCGTAA 701  
Qy 337 GGAGAACGAGTGCAGCTGGAAGAGCCATCAGTGTGCATCCACCCCTGAGGGCTGCTC 396  
Db 702 AGAAAAATCGGGGGCTGCTGAGAGTGCATTAATCTCTCTACTCTCTGAGGACCTC 761  
Qy 397 CTCGCTTGTAGATGATCTTGGAGATTAATGATTAAGAGGCTTAAGACACCAAGCGC 456  
Db 762 TCGGGCTTGTAGTCTTATTCGAGATTAATGATTAAGAGGCTTAAGATATAAAATTCAC 821  
Qy 457 TGACGAGTTCCTCGAGATCTCGGCCCAATAAATCTTGTAGGGCTCTCATTTGGCAA 516  
Db 822 AGAAGAGATCCCTTGAAGATTTAGCTCAATAAATCTTGTGACGCTCTATTGGTAA 881  
Qy 517 GGAAGGAGGAACTTGAAGAGGTTAGAGCAAGATACCGAGACAAAAATCACCATCTCCTC 576  
Db 882 AGAAGGAGAAATCTTAAAAAATTTAGCAAGACACACACTAAATCACGATATCTCC 941  
Qy 577 GTTGAAGACCTTACCTTTACACCTTGAGAGGACCATCACTGTGAAGGGGGCCATCGA 636  
Db 942 ATTGAGGAATTCAGCGCTGTATAATCCAGAACGCACTATTACAGTTAAAGCAATGTGA 1001  
Qy 637 GAATTTGTCAGGGCCGAGCAGAAATAATGAAGAACTTCGGAGGGCTATGAGAAATGA 696  
Db 1062 GACATGTGCCAAGCTGAGGAGGATCATGAGAAATCAGGAGTCTTATGAAATGA 1061  
Qy 697 TGTGCTGCGCATGA-----GCTCTACCTGATCCCTGGCTGAACTGGCTGCTGTAGG 750  
Db 1062 TATTGCTCTATGAATCTTCAAGCACATTTAATCTCTGATTAATCTGAACGCTTGGG 1121  
Qy 751 TCTTTTCCAGCTTCATCCAGCGAGTCCCGCGCTCCAGCAGCGTTACTTGGGGCTGC 810  
Db 1122 TCTGTTCCACCACTTCAGGATGCGAATCAGAAACGAGAGTGTTCATCTGTTTATCCACG 1181  
Qy 811 TCCCTATAGCTCTTTATGAGGCTCCGAGCAGGAGATGTTGAGGTTTATCCCGC 870  
Db 1182 TCTCCCTACCGCAGTTTGAACAATCAGAAACGAGAGTGTTCATCTGTTTATCCACG 1241  
Qy 871 CAGGAGTGGGGCCATCATCGGAGAGGGGAGCAGACATCAACAGCTCTCCGGTT 930  
Db 1242 TCTATCAGTGTGGCTCATCGGAGCAGGCGCAGCAGCATCAAGCAGCTTCTCGCTT 1301  
Qy 931 TGCCAGCGCTCCATCAAGATTGCAACCCGAAACACCTGACTCCAAAGTTTCTGATGGT 990  
Db 1302 TGCTGAGGTTCAATTAGATTGCTCCAGCGAAGCAGCATGCTTAAGTCAGGATGT 1361  
Qy 991 TATCATCATGAGCCCGCAGAGCCCAATTCAGGCTAGGGAAGAACTTATGGCAACT 1050  
Db 1362 GATTATCATGGAACCAACAGAGGCTCAGTTCAAGGCTCAGGGAAGAAATTTATGGAAAT 1421  
Qy 1051 CAAGGAGGAGAACTTCTTGTCCCAAGAGGAGTGAAGCTGGAGACCCACATACGTT 1110  
Db 1422 TAAAGAGAAACTTTGTTAGTCTTAAGAGAGGTGAACHTTGAAGTCTATCAGAGT 1481  
Qy 1111 GCAGCATCAGCAGCTGCGGGTCAATTTGCAAGGTGGAAGAAACGGTGAACGATTGCA 1170  
Db 1482 GCATCTTTGCTGTGCGAGAGTTATTGGAAGAGGAGGCAAAACGGTGAATGAATTC 1541  
Qy 1171 GAATTTGCGGAGCTGAGGTGGTAGTACCAAGAGCCAGCCCTGATGAGAACGACA 1230  
Db 1542 GAATTTGTCAAGTGCAGAGTGTGTTGCTCTCTGAGCAGACACCTGATGAGAAATGACA 1601  
Qy 1231 GGTTCGTGTAATCATCGGACATTTCTATGCGAGTCAAGTGGCTCAACGGAAGATCGG 1290  
Db 1602 AGTGGTTGTCAAATTAATCTGTCATCTTATGCTTGCAGGTTGCCAGAGAAAATTC 1661  
Qy 1291 AGATCTCTGGCCAGGTTAAGCAGCAGCATCAGAG 1327

Db 1662 GGAAATTTGACTAGTAAAGCAGCACCAACACAG 1698  
RESULT 10  
US-10-007-700-347  
; Sequence 347, Application US/10007700  
; Publication No. US2003064947A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Fanger, Neil  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Durham, Margarita  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Peckman, David W.  
; APPLICANT: Cai, Feng  
; APPLICANT: Foy, Teresa M.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.455C17  
; CURRENT APPLICATION NUMBER: US/10/007,700  
; CURRENT FILING DATE: 2001-11-30  
; NUMBER OF SEQ ID NOS: 469  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 347  
; LENGTH: 1740  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-007-700-347  
Query Match 34.1%; Score 582.2; DB 13; Length 1740;  
Best Local Similarity 66.2%; Pred. No. 6.9e-471;  
Matches 858; Conservative 0; Mismatches 433; Indels 6; Gaps 1;  
Qy 37 AGCCATCATGAAGTGAATGGCCACCACTTGGAGAACCATCGCTCGAAGGTCTCTACAT 96  
Db 402 AGACTAGACAACTGAATGATTTAGTTAGAGATTTTCACTTGAAGTAGCTATAT 461  
Qy 97 CCGCATGAGCAGATAGCAGGACCTGAGATGGCGCGGAGGGGCTTTGGCTCTCG 156  
Db 462 CCCTGATGAACGGCGCGCCAGCAAAACCCCTTTCAGCAGCCCGGAGGCTCGCGGGGCT 521  
Qy 157 GGGTCAGCCCGCCAGGGCTCACCTGTGGCAGCGGGGGCCAGCCAGCAGAGCAAGT 216  
Db 522 TGGCAGAGGGGCTCTCAAGGAGGGGTCTCCAGGATCCGTATCCNAGCAGAAACCATG 581  
Qy 217 GGACATCCCTCTTGGCTCTCTGGTGGCCACCCAGTATGTGGTGCCATTTATGGCAAGGA 276  
Db 582 TGAATTCCTCTGCGCTGCTGGTTCCACCCCAATTTGTTGGAGCCATCATAGAAAAGA 641  
Qy 277 GGGGGCCACCATCGGCAACATCAAAACAGACCCAGTCCAAAGATAGACCTGATAGGAA 336  
Db 642 AGGTGCCACCATTCGGAACATCACCAACAGACCCAGTCTAAATTCGATGTCCACCGTAA 701  
Qy 337 GGAGAACGAGTGCAGCTGGAAGAGCCATCAGTGTGCATCCACCCCTGAGGGCTGCTC 396  
Db 702 AGAAAAATCGGGGGCTGCTGAGAGTGCATTAATCTCTCTACTCTCTGAGGACCTC 761  
Qy 397 CTCGCTTGTAGATGATCTTGGAGATTAATGATTAAGAGGCTTAAGACACCAAGCGC 456  
Db 762 TCGGGCTTGTAGTCTTATTCGAGATTAATGATTAAGAGGCTTAAGATATAAAATTCAC 821  
Qy 457 TGACGAGTTCCTCGAGATCTCGGCCCAATAAATCTTGTAGGGCTCTCATTTGGCAA 516  
Db 822 AGAAGAGATCCCTTGAAGATTTAGCTCAATAAATCTTGTGACGCTCTATTGGTAA 881

QY 517 GGAAGGACGACCTGAAGAGGTAGAGCAAGATACCGAGACAAATAATCACCATCTCTC 576  
Db 882 AGAGGAAGAAATCTTAAATAAATGAGCAAGACAGACACTAAATACAGATATCTCC 941  
QY 577 GTTGAAGACCTTACCCCTTTAAACCCCTGAGAGGACCATCACTGTGAAGGGGGCCATCGA 636  
Db 942 ATTGCAAGAAATGACGCTGTATAATCCAGAACGCACTATTACAGTTAAAGGCAATGTTGA 1001  
QY 637 GAATTTGTCAGGGCCGAGCAGCAAAATATAGAAAGTTTCGGAGGCTTATGAGATGA 696  
Db 1002 GACATGTGCAAAAGTGTAGAGGAGATCATGAAGAAATCAAGGAGTCTTATGAAATGA 1061  
QY 697 TGTGGCTGCATGA-----GCTCTCACCTGATCCCTGCAACCTGGCTGTGTAGG 750  
Db 1062 TATTGCTTCTATGAATCTTCAAGCACATTTAATCTCTGATTAATCTGAAGCCTTGGG 1121  
QY 751 TCTTTTCCAGCTTCATCCAGCGCAGTCCCGCGCTCCCGAGCAGCGTTACTGGGGTGC 810  
Db 1122 TCTGTTCCACCACTTTCAGGGATGCCACCTCCACCTCAGGGCCCCCTTCAGCCATGAC 1181  
QY 811 TCCCTATAGCTCTTATGCGAGCTCCCGAGCAGGAGATGTCAGGTGTTTATCCCGC 870  
Db 1182 TCTGCTTACCGCGAGTTTGCAATCAGAAACGGAGACTGTTTCATCTGTTTATCCCGC 1241  
QY 871 CCAGGACGTGGCGGCATCATCGGCAAGAGGGGAGCAGACATCAAAACAGCTCTCCCGTT 930  
Db 1242 TCTATCAGTCGGTGCATCATCGGCAAGCAGGGCCAGCACATCAAGCAGCTTCTCGCTT 1301  
QY 931 TGCCAGCGCTCCATCAAGATTCACACCGACCGAAACACTCACTCAAGTTCGTATGGT 990  
Db 1302 TCTCGAGCTTCAATTAAGATTGCTCCAGCGGAAGCACCAGATGCTAAAGTGAGGATGGT 1361  
QY 991 TATCATCTAGTCAGCGCCAGAGGCCAAATCAAGGCTCAGGGAAGAAATCTATGCAAACT 1050  
Db 1362 GATTATCACTGGACCAACAGAGGCTCAGTTCAAGGCTCAGGGAAGAAATTTATGAAAT 1421  
QY 1051 CAGAGAGAGAACTTCTTTGGTCCCAAGAGGAAGTGAAGTGGAGACCCACATACGTGT 1110  
Db 1422 TAAAGAGAAATCTTGTAGTCTTAAAGAGAGGTGAATCTTGAAGCTCATATCAGAGT 1481  
QY 1111 GCAGCATCAGCAGCTGCGCGGTTCATTGGCAAGAGTGGAAACCGTGAACGAGTTGCA 1170  
Db 1482 GCCATCTTGTGCTGCGAGATTATGGAAAGAGGCAAAACGGTGAATGAATTC 1541  
QY 1171 GAATTTGACGAGCTGAGGTGGTAGTACCAAGAGACAGACCCCTGATGAGACGACCA 1230  
Db 1542 GAAATTTGTCAAGTGCAGAAATGTTGTCCCTCGTGAACAGACACTGATGAGAAATGACCA 1601  
QY 1231 GGTATCTGGAATATCATCGACATTTCTATGCGAGTCAAGTGGCTCAAGGAGATCGG 1290  
Db 1602 AGTGGTGTCAAAATTAATGCTGACTTCTATGCTTGCAGGTGCCAGAGTCCCGAGAAATTC 1661  
QY 1291 AGACATCTGCGCCAGGTTAAGCAGCAGCATCAGAA 1327  
Db 1662 GGAAATCTGACTCAGGTAAAGCAGCAGCAACACAG 1698

## RESULT 11

US-10-117-982-347  
; Sequence 347, Application US/10117982  
; Publication No. US20030138438A1  
; GENERAL INFORMATION:  
; APPLICANT: Foy, Teresa M.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Vedrick, Thomas S.  
; APPLICANT: Carter, Derrick  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Mericle, Barbara  
; APPLICANT: Spies, Gregory A.  
; APPLICANT: Fan, Liqun

; APPLICANT: Wang, Tongtong  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455G18  
; CURRENT APPLICATION NUMBER: US/10/117,982  
; CURRENT FILING DATE: 2002-04-05  
; NUMBER OF SEQ ID NOS: 484  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 347  
; LENGTH: 1740  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-117-982-347

Query Match 34.1%; Score 582.2; DB 15; Length 1740;  
Best Local Similarity 66.2%; Pred. No. 6.9e-171;  
Matches 858; Conservative 0; Mismatches 433; Indels 6; Gaps 1;

QY 37 AGCCATCATGAAGCTGAATGGCCACCAGTTTGAGAACCATGCGCTGAAGGTCTCTACAT 96  
Db 402 AGCCTAGACAACTGAATGGATTTTCAGTTAGAGAAATTTACCTTGAAGTAGCCTATAT 461  
QY 97 CCCGATGAGCAGATAGCAGGGACCTGAGAATGGCGCCGAGGGGCTTTGGCTTCG 156  
Db 462 CCTGATGAAGCGCGCCCGCAGCAAAACCCCTTGAGCAGACCCCGAGGTGCGCGGGCT 521  
QY 157 GGTGACGCCCCCAGGGCTCACCTGTGCGAGGGGGCCCCAGCCAGCAGCAGCAAGT 216  
Db 522 TGGCAGAGGGGCTCCTCAAGCAGGGGTCTCCAGATCCGTATCCAGACGAAACCATG 581  
QY 217 GGCATATCCCTTGGCTCTCTGCTGCTCCACCCAGTATGTGGGTGCATTTATGGCAAGA 276  
Db 582 TGATTTGCTCTGCGCTGCTGTTTCCCAACCAATTTGTTGGAGCCATCATAGGAAAGA 641  
QY 277 GGGGCGCAGCATCCGACATCACAACAGACAGCCAGTCCAGATAGAGTAGCATAGAA 336  
Db 642 AGTGCCACCATTCGGAACATCACCAGACAGACCCAGTCTAAATTCGATGTCCACCGTAA 701  
QY 337 GGAGAACGAGCTGAGCTGAAAGAACCATCAGTGTGCACTCCACCTCGAGGGGTGCTC 396  
Db 702 AGAAATGGGGGCTGCTGAGAGTCCGATTAATCTCTCTCTCTCTCTCTCTCTCTCTCT 761  
QY 397 CTCGCTTTGAAGATGATCTTGGAGATTATGCATTAAGAGGCTAAGACACCAAAACGCG 456  
Db 762 TCGGGCTTTGAAGTCTATTTCTGGAGATTATGCATTAAGAGGCTCAAGATATAAAATTCAC 821  
QY 457 TGACGAGGTTCCTCTGAAGATCTCGGCCATATAAATTTGTAGGGCTCTCATTTGGCAA 516  
Db 822 AGAAGATCTCCCTTGAAGATTTTAGCTCATATAAATTTGTTGGAGCTCTTTATGGTAA 881  
QY 517 GGAAGAGGAACTGAAAGAGGTAGACAAAGATACCGAGACAAATAATCACCATCTCTC 576  
Db 882 AGAAGAAATATCTTAAATAAATTTGAGCAAGACAGACACTAAATCACGATATCTCC 941  
QY 577 GTTGAAGACCTTACCTTTTACACCTTGAAGAGGACCATCACTGTGAAGGGGCCATCGA 636  
Db 942 ATTGAGGAATTCAGCTGTATATCCAGAACGCACTATTACAGTTAAAGGCAATGTTGA 1001  
QY 637 GAATTTGTCAGGGCCGAGCAGGAAATTAATGAAGAAAGTTCGGGAGGCTTATGAGATGA 696  
Db 1002 GACATGTGCAAAAGTGTAGAGGAGATCATGAAGAAATCAGGAGTCTTATGAAATGA 1061  
QY 697 TGTGCTGCATGA-----GCTCTCACCTGATCCCTGCAACCTGGCTGTGTAGG 750  
Db 1062 TATTGCTTCTATGAATCTTCAAGCACATTTAATCTCTGATTAATCTGAAGCCTTGGG 1121  
QY 751 TCTTTTCCAGCTTCATCCAGCGAGTCCCGCGCTCCCGAGCAGCGTACTTGGGGTGC 810  
Db 1122 TCTGTTCCACCACTTTCAGGGATGCCACCTTCCACCTCAGGGCCCCCTTCAGCCATGAC 1181  
QY 811 TCCCTATAGCTCTTATGCGAGCTCCCGAGCAGGAGATGTCAGGTGTTTATCCCGC 870  
Db 1182 TCTCTCCCTACCGCGAGTTTGCAATCAGAAACGGAGACTGTTTCATCTGTTTATCCCGC 1241



QY 871 CCAGGAGTGGCGCCATCATCGGCAAGAAGGGGAGACATCAAAACAGCTCTCCGGGT 930  
DB 1242 TCTATCAGTCGGTGCATCATCGGCAAGCAGGGGAGACATCAAGCAGCTTCTCGCT 1301  
QY 931 TGCCAGGCGCTCCATCAAGATTGACACCCGMAACCTGACTCCAAAGTTCGTATGTT 990  
DB 1302 TGCTGAGCTCAATTAAGATTGCTCAGCGAAGCACCAGATGCTAAAGTGAGATGGT 1361  
QY 991 TATCATCTAGGACCGCCAGAGGCCAAATCAAGGCTCAGGGAAGAATCTATGGCAAACT 1050  
DB 1362 GATTATCATCTGACCAACAGAGGCTCAGTTCAAGGCTCAGGGAAGAATTTATGGAATAAT 1421  
QY 1051 CAAGGAGAGAACTCTTTGGTCCCAAGAGGAAGTGAAGCTGAGACCCACATAGTGT 1110  
DB 1422 TAAAGAAGAACTTTGTTAGTCTCTAAAGAGAGGTGAAGCTTGAAGCTCATATCAGAGT 1481  
QY 1111 GCCAGCATCAGCAGTGGCGCGGTCAATTTGGCAAGGTGGAACACGGTGAACGAGTTGCA 1170  
DB 1482 GCCATCCTTTGCTGCTGCAGAGTTATTGGAAGAGGAGGCAACGGTGAATGACTTCA 1541  
QY 1171 GAATTTGAGGAGCTGAGGTGGTAGTACCAAGAGACAGACCCCTGATGAGAACGACCA 1230  
DB 1542 GAATTTGTCAGAGTGCAGAGTGTGTCCCTCGTGACAGACACCTGATGAGAAATGACCA 1601  
QY 1231 GGTCACTGTGAATCATCTGACATTTCTATGCCAGTCAGATGCTCAACGGAAGATCG 1290  
DB 1602 AGTGGTTGCAAAATCACTGCTACTTCTATGCTTCCAGGTTGCCAGAGAAAATTC 1661  
QY 1291 AGACATCTGCCCGAGTTAAGCAGCAGCATCAGAAG 1327  
DB 1662 GGAAATTCGACTCAGGTAAAGCAGCACCACCAACAG 1698

## RESULT 12

US-10-117-982-478  
; Sequence 478, Application US/10117982  
; Publication No. US20030138438A1  
; GENERAL INFORMATION:  
; APPLICANT: Foy, Teresa M.  
; APPLICANT: Farger, Gary R.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Darriek  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Mericle, Barbara  
; APPLICANT: Spies, Gregory A.  
; APPLICANT: Fan, Liqun  
; APPLICANT: Wang, Tongtong  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.455C18  
; CURRENT APPLICATION NUMBER: US/10/117,982  
; CURRENT FILING DATE: 2002-04-05  
; NUMBER OF SEQ ID NOS: 484  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 478  
; LENGTH: 1740  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-117-982-478

Query Match 34.1%; Score 582.2; DB 15; Length 1740;  
Best Local Similarity 66.2%; Pred. No. 6.9e-171;  
Matches 858; Conservative 0; Mismatches 433; Indels 6; Gaps 1;  
QY 37 AGCCATCATGAGCTGATGCCACCAAGTTGGAGAACCATGCCCTGAGCTCTCCATCAT 96  
DB 402 AGCACTAGACAACTGAATGATTTCAGTTAGAGAAATTTTCACCTTGAAGTAGCCTATAT 461  
QY 97 CCCCGATGAGCAGATAGCAGAGGACCTGAGAAATGGCGCCGAGGGGGCTTTGGCTCTCG 156

DB 462 CCCTGATGAAACGGCCGCCCCAGCAAAACCCCTTTCAGCAGACCCCGAGGTTCGCGGGGCT 521  
QY 157 GGGTCAGCCCGCCAGGGCTCACTCTGTGCGAGGGGGGCCAGCAAGCAGCAGCAAGT 216  
DB 522 TGGGCGAGGGGCTCCTCAAGCAGGGGTCTCAGGATCCGTATCCAGAGAAACCATG 581  
QY 217 GGCATCTCCCTTCGGCTCCTGGTGCACCCAGTATGCGGTGCCATTTATGGCAAGA 276  
DB 582 TGATTTGGCTCTGGGCTGCTGGTTCCCAACCCAAATTTGTGGAGCCATCATAGGAAAAGA 641  
QY 277 GGGGGCCACCATCCGCAACATCAAAACAGACCCAGTCCAGATAGAGTGCATAGGAA 336  
DB 642 AGTGGCCACATTCGGAACATCACAAACAGACCCAGTCTAAATCGATGTCACCGTAA 701  
QY 337 GGAAGACGAGGTGACGCTGAAAAAGCCATCAGTGTGCACTCCACCCCTGAGGGCTGCTC 396  
DB 702 AGAAATGCGGGGCTGCTGAGAAGTCGATTACTATCTCTCTACTCTCTGAGGACCTC 761  
QY 397 CTCGGCTTGAAGATGATCTTGGAGATTATGATAAAGAGGTGAGGACCAACAAACGGC 456  
DB 762 TCGGGCTTGTAAAGTCTTATTTCTGGAGATTATGATAAAGAGGTGAGGATATAAAATTC 821  
QY 457 TGACGAGGTTCCTCCCTGAAGATCCTGGCCCATATAAATCTTTGTAGGCGCTCTCATTTGGCAA 516  
DB 822 AGAAGATCCCTTGAAGATTTTAGCTCATATAACTTTTGTGGACGCTTTATTGGTAA 881  
QY 517 GGAAGACGGAACCTGGAAGAGGTAGCAAGATACCGAGACAAAAATCACCATCTCTC 576  
DB 882 AGAAGAGAAAGAAATCTTAAAAAAATTTGAGCAAGACACAGACACTAAAAATCAGTATCTCC 941  
QY 577 GTTGAAGACCTTACCTTTTCAACCTGAGAGGACCATCAGTGTGAAGGGGGCCATCGA 636  
DB 942 ATTGCAGGAATTGACGCTGTATAATCCAGAACGCACTATTACAGTTAAAGGCAATGTTGA 1001  
QY 637 GAATTTTCAGGGCGGAGCAGGAAATAATGAAGAAAGTTCCGGAGGCGCTATGAGAATGA 696  
DB 1002 GACATGTGCAAAAGCTGAGGAGGAGATCATGAAGAAATCAGGGAGTCTTTATGAAATGA 1061  
QY 697 TGTGGCTGCCATGA-----GCTCTCACCTGATCCCTGSCCTGAACCTGGCTGCTGTAGG 750  
DB 1062 TATTCCTTCTATGAATCTTTCAAGCACATTTAATCTCGGATTTAAATCTGAACGCTTGGG 1121  
QY 751 TCTTTTCCAGCTTTCATCCAGGCGCAGTCCCGCGGCTCCAGCAGCGTTTACTGGGCGTGC 810  
DB 1122 TCTGTTCCCAACCACTTCAGGGATGCCACTTCCCACTCAGGGCCCCCTTCAGCCATGAC 1181  
QY 811 TCCCTATAGCTCCTTTATGCAAGCTCCCGAGCAGGAGATGGTGCAGGTGTTTATCCCCGC 870  
DB 1182 TCTTCCCTACCCGCGAGTTTGACAAATCAGAAACGGAGACTGTTTCATCTTTTATCCAGC 1241  
QY 871 CCAGGAGTGGCGGCATCATCGCAAGAAAGGGGAGCAGACATCAACACAGCTCTCCCGTT 930  
DB 1242 TCTATCAGTCGGTGCATCATCGCAAGCAGGGGCGAGACATCAAGAGCTTTTCTCGTT 1301  
QY 931 TGCCAGGCGCTCCATCAAGATTGCAACCCGAAACACCTGACTCCAAAGTTTCGTATGTT 990  
DB 1302 TGTGAGCTTCAATTAAGATTGCTCCAGCGGAAGCACCAGATGCTAAAGTCAAGATGTT 1361  
QY 991 TATCATCTGACCGCAGAGGCCAATTCAGGCTCAGGGAAGAAATCTATGGCAACT 1050  
DB 1362 GATTATCACTGGACCAACAGAGGCTCAGTTCAAGGCTCAGGGAAGAAATTTATGGAATAAT 1421  
QY 1051 CAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAGTGAAGCTGGAGACCCACATACGTTT 1110  
DB 1422 TAAAGAGAAACTTTGTTAGTCTTAAAGAGAGGTGAACCTTGAAGCTCATATCAGAT 1481  
QY 1111 GCACGATCAGCAGCTGGCCGGGTCAATTTGGCAAGAGGTGGAAGAAACCGTGAACAGTTGCA 1170  
DB 1482 GCACTCCTTTGCTGTCGAGAGTTTATTTGAAAAAGAGGCAAAACCGTGAATGAATCTCA 1541  
QY 1171 GAATTTGAGCGGAGCTGAGGTGGTAGTACCAGAGACCAAGCCCTGATGAGAACGACCA 1230  
DB 1542 GAATTTGTCAAGTGAGAGTTGTTGTCCTCGTCCAGACACCTGATGAGAAATGACCA 1601

QY 1231 GGTCACTGTAATCATCGGACATTTCTATGCGAGTCAGATGGCTCAACGGAGATCCG 1290  
Db 1602 AGTGGTGTGTAATACTGCTCACTTCTATGCTTGCACAGTTGCCAGAGAAATTTCA 1661

QY 1291 AGACATCTGCCAGGTTAAGCAGCAGCATCAAG 1327  
Db 1662 GGAATTTCTACTCAGGTAAAGCAGCACAACACAG 1698

RESULT 13  
US-10-313-986-347  
; Sequence 347, Application US/10313986  
; Publication No. US20030236209A1  
; GENERAL INFORMATION:  
; APPLICANT: Foy, Teresa M.  
; APPLICANT: McNabb, Andria  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Wang, Tongtong  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.455C19  
; CURRENT APPLICATION NUMBER: US/10/313.986  
; NUMBER OF SEQ ID NOS: 560  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 347  
; LENGTH: 1740  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-313-986-347

Query Match 34.1%; Score 582.2; DB 16; Length 1740;  
Best Local Similarity 66.2%; Pred. No. 6.9e-171;  
Matches 858; Conservative 0; Mismatches 433; Indels 6; Gaps 1;

QY 37 AGCCATCATGAAGCTGAATGCCACCACTGGAGAACCATGCCCTGAAGCTCTCTACAT 96  
Db 402 AGCACTAGACAACTGAATGATTTTCAGTTAGAGAAATTTACCTTTGAAGTAGCTATAT 461

QY 97 CCCCAGCAGCAGATAGCACAGGACCTTGAAATGGCGCGGAGGGGCTTTGGCTCTCG 156  
Db 462 CCTGATGAACGGCGCCGAGCAAAACCCCTTGCAGCAGCGCCGAGGTGCGCGGGGCT 521

QY 157 GGGTCAGCCCCGACGGGCTCACTGTGGAGCGGGGGCCCCAGCAGCAGCAAGT 216  
Db 522 TGGCAGAGGGGCTCTCAAGGCGAGGGTCTCCAGGATCCGATCCAAAGCAGAAACCATG 581

QY 217 GGACATCCCTTCGGCTCTGGTGGCCACCCAGTATGTGGTGCCCATTTTGGCAAGGA 276  
Db 582 TGAATTTGCTCTGGCTCTGGTGGTCCACCCCAATTTTGGAGCCATCATAGAAAGA 641

QY 277 GGGGCCACCATCCGCAACATCAAAACAGACCCAGTCCAAAGATAGACGTGCATAGGAA 336  
Db 642 AGGTGCCACCATTCGGAACATCAACAAAGACAGCCAGTCTAAATTCGATGTCCACCGTAA 701

QY 337 GGGAACCCAGGTGCAGCTGAAAAGCCATCAGTGTGCCTCCACCCCTGAGGCTGCTC 396  
Db 702 AGAATATCGGGGCTGTGAGAGTGTGATTTACTATCTCTCTACTCTGAGGACCTC 761

QY 397 CTCGCTTGTGAAGATGATCTTTGGAGATATGATCAAAAGAGGCTAAGGACCAAAACGGC 456  
Db 762 TCGGCTTTGAAGTCTATTTCTGGAGATATGATCAAAAGAGGCTCAAGATATAAAATTCAC 821

QY 457 TCACAGGCTTCCCTGAGATCCTGGCCCAATAAATTTGTAGGGGTCTCATTTGGCAA 516  
Db 822 AGAAGAGATCCCTTGAAGATTTTAGCTCAATAAATTTTGTGAGGCTCTTATTGTAA 881

QY 517 GGAAGCGGAACTTGAAGAGAGGTAGACCAAGATACCGAGACAAATAATCACTCTCCTC 576  
Db 882 AGAAGGAAGAAATCTTTAAAAAAATTTGAGCAAGACAGACACACTAAATTCAGTATCTCC 941

QY 577 GTTGAAGACCTTAACCTTTTACAAACCTGTAGAGAGACCATCACTGTGAAGGGGGCCATCGA 636  
Db 942 ATTGCAGGAATTTGACGCTGTATAATCCAGAAACGCACTATTACAGTTAAAGGCAATGTGA 1001

QY 637 GAAATTTTGGAGGCGGAGCAGGAATAATGAAGAAAGTTCCGGAGGCGCTATGAGAATGA 696  
Db 1002 GACATGTGCCAAAGCTGAGGAGAGATCATGAAGAAATCAGGAGTCTTTATGAATAATGA 1061

QY 697 TGTGGCTGCCATGA-----GCTCTCACCTGATCCCTGGCTGAACCTGGCTGTGTAGG 750  
Db 1062 TATTGCTTCTATGAATCTTCAAGCACATTTAATTCCTGGATTAATCTGAACGCTTTGGG 1121

QY 751 TCTTTTCCAGGTTTCATCCAGGAGTCCCGCGCTCCAGCAGCGTTTACTGGGGCTGC 810  
Db 1122 TCTGTTCCACCCACTTCAGGGATGCCACCTCCCACTCAGGGCCCTCTCAGGCATGAC 1181

QY 811 TCCCTATAGCTCTTATGTCAGGCTCCGAGCAGAGATGGTCAGAGTGTATATCCCCGC 870  
Db 1182 TCTCCCTACCGGAGTTTGAGCAATCAGAAACGAGACTGTTCATCTGTTTATCCCCAGC 1241

QY 871 CCAGCAGTGGGCGCCATCATCGGCAAGAGGGGAGCACAATCAACAGCTCTCCCGTT 930  
Db 1242 TCTATCAGTCGGTGCATCATCGCAGCAGGCGCAGCACATCAAGCAGCTTTCTCGCT 1301

QY 931 TGCCAGCGCTCCATCAAGATTGCACACCCGAAACCTGACTCCAAAGTTCTGTATGGT 990  
Db 1302 TGTGGAGCTTCAATTAAGATTTCTCCAGCGAAGCAGATGTTAAGTGAGGATGGT 1361

QY 991 TATCATCACTGGACCGCCAGAGGCCCAATTCAAGGCTCAGGGAAGAAATCTATGCAAACT 1050  
Db 1362 GATTATCACTGGACCAACAGAGGCTCAGTTCAAGGCTCAGGGAAGAAATTTATGAAAAAT 1421

QY 1051 CAGCAGAGAGACTTCTTTGCTCCAGAGGAGTGAAGCTGGAGACCCACATACGTGT 1110  
Db 1422 TAAAGAAGAAACTTTGTTAGTCTTAAAGAAGAGTGAAACTTGAAGCTCATTCAGAGT 1481

QY 1111 GCCAGCATCAGCAGCTGGCGGCTCATTTGGCAAAAGGTGGAAAAACGGTGAACGAGTTGCA 1170  
Db 1482 GCCATCTTCTGCTGTCAGAGATTTTGGAAAAAGGAGGCAAAACGGTGAATGAATTC 1541

QY 1171 GAATTTGAGCAGCTGAGTGTGTAGTACCAAGACAGCAGACCCCTGATGGAACGACCA 1230  
Db 1542 GAAATTTGCAAGTGCAGAAAGTTGTCTCCCTCGTGACAGACACCTGATGAGAATGACCA 1601

QY 1231 GGTATCTGTGAAAAATCATCGGACATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCCG 1290  
Db 1602 AGTGGTTGTCAAAATACTGGTCACTTCTATGCTTGCCAGGTTGCCAGAGAAAAATTC 1661

QY 1291 AGACATCTCGGCCAGGTTAAGCAGCAGCATCAGAA 1327  
Db 1662 GGAATTTCTACTCAGGTAAAGCAGCACCACCAACAG 1698

RESULT 14  
US-10-313-986-478  
; Sequence 478, Application US/10313986  
; Publication No. US20030236209A1  
; GENERAL INFORMATION:  
; APPLICANT: Foy, Teresa M.  
; APPLICANT: McNabb, Andria  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Wang, Tongtong  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.455C19  
; CURRENT APPLICATION NUMBER: US/10/313.986  
; NUMBER OF SEQ ID NOS: 560  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 478  
; LENGTH: 1740  
; TYPE: DNA



```

; ORGANISM: Homo sapiens
US-10-313-986-478

Query Match      34.1%; Score 582.2; DB 16; Length 1740;
Best Local Similarity 66.2%; Pred. No. 6.9e-171;
Matches 858; Conservative 0; Mismatches 433; Indels 6; Gaps 1;

QY 37 AGCCATCATGAAGCTGAATGCCACCGAGTTCGAGAACCATCGCTCGAAGCTCTCTACAT 96
DB 402 AGCACTAGACAACTGAATGATTTAGTTAGAGATTTACCTTGAAAGTAGCCTATAT 461
QY 97 CCCGATGAGCAGATAGACAGGACCTTGAGAAATGGGCGCGAGGGGGCTTTGGCTCTCG 156
DB 462 CCTGATGAAACGGCGCGCCACCAAAACCCCTTGACAGAGCCCGGAGGTTCGCGGGGT 521
QY 157 GGGTCAGCCCCCGCAGGGCTCACCTGTGGCAGCGGGGGCCCGAGCCAGCAGCAGCAAGT 216
DB 522 TGGGCGAGGGGGCTCCTCAAGGCAGGGGTCTCCAGGATCCGATATCAAGAGCAAAACCATG 581
QY 217 GGACATCCCCCTTCGGCTCCCTGGTCCACCCAGTATGTGGGTGCCATTATTGGCAAGA 276
DB 582 TGATTTGCTCTGCGCTTCTGGTTCCCAACCCATTTGTTGGAGCCATCATAGGAAAGA 641
QY 277 GGGGCGCCACCATCGCAACATCACAAGACAGACCCAGTCCAGATAGAGTGCATAGGAA 336
DB 642 AGGTGCCACCAATCGGAACATCACCAAAACAGACCCAGTCTAAATCGATGTCCACCGTAA 701
QY 337 GGAGAACGACGCTGAGCTGAAAGGACCATCAGTGTGCACCTCCACCCCTGAGGGCTGCTC 396
DB 702 AGAAATGCGGGGCTGCTGAGATGCAATTAATCTCTCTACTCTGAGGACCTC 761
QY 397 CTCGCTTGTAAGATGATCTTGGAGATATGCAATAAGAGGCTTAAGACACCAAAACGGC 456
DB 762 TCGGGCTTGTAAGTCTATCTTGGAGATATGCAATAAGAGGCTCAAGATATAAAATTCAC 821
QY 457 TGACGAGTTCCTCCCTGAGATCTCTGGCCCAATAAATCTTGTAGGGGCTCTATTGSCAA 516
DB 822 AGAAGATATCCCTTGAAGATTTAGCTCATATAAATCTTGTGAGCGTCTATTGGTAA 881
QY 517 GGAAGGACGAACTGAAAGAGGTAGAGCAAGATACCGAGACAAAAATCACCATCTCTCCTC 576
DB 882 AGAAGGAAGAAATCTTAAAAAATGAGCAAGACACAGACACTAAATACGATATCTCC 941
QY 577 GTTCCAGACCTTACCTTTACACCTTGAGAGGACCATCACTGTGAAGGGGGCATCGA 636
DB 942 ATTGCAGAAATGACGTGTATAATCCAGAACGCACTATTACAGTTAAAGGCAATGTTGA 1001
QY 637 GAATTTGTGAGGGCCGAGCAGGAAATAATGAAGAAAGTTCGGAGGCTTATGAGATGA 696
DB 1002 GACATGTCCAAAGCTGAGGAGGAGATCATGAAGAAATCAGGAGTCTTATGAATGA 1061
QY 697 TGTGGCTGCATGA-----GCTCTACCTGATCCCTGGCTGAACCTGCTGCTGTGAGG 750
DB 1062 TATTGCTTCTATGAATCTTCAAGCACATTTAAATCTCTGGATTAATCTGAACGCTTGGG 1121
QY 751 TCTTTTCCAGCTTCATCCAGCGAGTCCCGCGCTCCAGCAGGCTTACTGGGCTGC 810
DB 1122 TCTGTTCCACCACTTACGGATGCGACCTTCCACCTCAGCGCCCTTCAGCCATGAC 1181
QY 811 TCCCTATAGTCTTATATGAGGCTCCCGAGCAGGAGATGTGTCAGGTGTTTATCCCGC 870
DB 1182 TCCTCCCTACCGCAGTTTGAGCAATCAGAAACGAGACTGTTTCATCTGTTTATCCAGC 1241
QY 871 CCAGGCGATGGCGGCATCATCGCAAGAGGGGCGAGCACAATCAACAGCTCTCCCGGTT 930
DB 1242 TCTATAGTGGTGGCCATCATCGCAAGAGGGGCGAGCACAATCAAGCAGCTTCTCGCTT 1301
QY 931 TGGCAGGGCTTCCATCAAGATTGACCAACCGCAACACCTGACTCCCAAGTTCGTATGGT 990
DB 1302 TGCTGAGCTTCAATTAGATTGCTCCAGCGGAGCAACGAGTCTTAAGTGGAGTGGT 1361
QY 991 TATCATCATGGACCGCGAGGGCCCAATTCAGGGCTCAGGGAGAAATCTATGGCAAACT 1050
DB 1361 TATCATCATGGACCGCGAGGGCCCAATTCAGGGCTCAGGGAGAAATCTATGGCAAACT 1050

RESULT 15
US-09-897-778-447
; Sequence 447, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C16
; CURRENT APPLICATION NUMBER: US/09/897,778
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 447
; LENGTH: 1743
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-897-778-447

Query Match      34.1%; Score 582.2; DB 9; Length 1743;
Best Local Similarity 66.2%; Pred. No. 6.9e-171;
Matches 858; Conservative 0; Mismatches 433; Indels 6; Gaps 1;

QY 37 AGCCATCATGAAGCTGAATGCCACCGAGTTCGAGAACCATCGCTCGAAGCTCTCTACAT 96
DB 402 AGCACTAGACAACTGAATGATTTAGTTAGAGATTTACCTTGAAAGTAGCCTATAT 461
QY 97 CCCGATGAGCAGATAGACAGGACCTTGAGAAATGGGCGCGAGGGGGCTTTGGCTCTCG 156
DB 462 CCTGATGAAACGGCGCGCCACCAAAACCCCTTGACAGAGCCCGGAGGTTCGCGGGGT 521
QY 157 GGGTCAGCCCCCGCAGGGCTCACCTGTGGCAGCGGGGGCCCGAGCCAGCAGCAGCAAGT 216
DB 522 TGGGCGAGGGGGCTCCTCAAGGCAGGGGTCTCCAGGATCCGATATCAAGAGCAAAACCATG 581
QY 217 GGACATCCCCCTTCGGCTCCCTGGTCCACCCAGTATGTGGGTGCCATTATTGGCAAGA 276
DB 582 TGATTTGCTCTGCGCTTCTGGTTCCCAACCCATTTGTTGGAGCCATCATAGGAAAGA 641
QY 277 GGGGCGCCACCATCGCAACATCACAAGACAGACCCAGTCCAGATAGAGTGCATAGGAA 336
DB 642 AGGTGCCACCAATCGGAACATCACCAAAACAGACCCAGTCTAAATCGATGTCCACCGTAA 701
QY 337 GGAGAACGACGCTGAGCTGAAAGGACCATCAGTGTGCACCTCCACCCCTGAGGGCTGCTC 396
DB 702 AGAAATGCGGGGCTGCTGAGATGCAATTAATCTCTCTACTCTGAGGACCTC 761
QY 397 CTCGCTTGTAAGATGATCTTGGAGATATGCAATAAGAGGCTTAAGACACCAAAACGGC 456
DB 762 TCGGGCTTGTAAGTCTATCTTGGAGATATGCAATAAGAGGCTCAAGATATAAAATTCAC 821
QY 457 TGACGAGTTCCTCCCTGAGATCTCTGGCCCAATAAATCTTGTAGGGGCTCTATTGSCAA 516
DB 822 AGAAGATATCCCTTGAAGATTTAGCTCATATAAATCTTGTGAGCGTCTATTGGTAA 881
QY 517 GGAAGGACGAACTGAAAGAGGTAGAGCAAGATACCGAGACAAAAATCACCATCTCTCCTC 576
DB 882 AGAAGGAAGAAATCTTAAAAAATGAGCAAGACACAGACACTAAATACGATATCTCC 941
QY 577 GTTCCAGACCTTACCTTTACACCTTGAGAGGACCATCACTGTGAAGGGGGCATCGA 636
DB 942 ATTGCAGAAATGACGTGTATAATCCAGAACGCACTATTACAGTTAAAGGCAATGTTGA 1001
QY 637 GAATTTGTGAGGGCCGAGCAGGAAATAATGAAGAAAGTTCGGAGGCTTATGAGATGA 696
DB 1002 GACATGTCCAAAGCTGAGGAGGAGATCATGAAGAAATCAGGAGTCTTATGAATGA 1061
QY 697 TGTGGCTGCATGA-----GCTCTACCTGATCCCTGGCTGAACCTGCTGCTGTGAGG 750
DB 1062 TATTGCTTCTATGAATCTTCAAGCACATTTAAATCTCTGGATTAATCTGAACGCTTGGG 1121
QY 751 TCTTTTCCAGCTTCATCCAGCGAGTCCCGCGCTCCAGCAGGCTTACTGGGCTGC 810
DB 1122 TCTGTTCCACCACTTACGGATGCGACCTTCCACCTCAGCGCCCTTCAGCCATGAC 1181
QY 811 TCCCTATAGTCTTATATGAGGCTCCCGAGCAGGAGATGTGTCAGGTGTTTATCCCGC 870
DB 1182 TCCTCCCTACCGCAGTTTGAGCAATCAGAAACGAGACTGTTTCATCTGTTTATCCAGC 1241
QY 871 CCAGGCGATGGCGGCATCATCGCAAGAGGGGCGAGCACAATCAACAGCTCTCCCGGTT 930
DB 1242 TCTATAGTGGTGGCCATCATCGCAAGAGGGGCGAGCACAATCAAGCAGCTTCTCGCTT 1301
QY 931 TGGCAGGGCTTCCATCAAGATTGACCAACCGCAACACCTGACTCCCAAGTTCGTATGGT 990
DB 1302 TGCTGAGCTTCAATTAGATTGCTCCAGCGGAGCAACGAGTCTTAAGTGGAGTGGT 1361
QY 991 TATCATCATGGACCGCGAGGGCCCAATTCAGGGCTCAGGGAGAAATCTATGGCAAACT 1050
DB 1361 TATCATCATGGACCGCGAGGGCCCAATTCAGGGCTCAGGGAGAAATCTATGGCAAACT 1050
```

QY 337 GGAGACGCGAGTCCAGTGAARAAGCCATCAGTGTGCACTCCACCCCTGAGGGCTGCTC 396  
Db |||||  
QY 702 AGAAAAATGCGGGGCTGCTGAGAGTCAATTAATCTTCTACTCTCTGAGGACCTC 761  
Db |||||  
QY 397 CTCGGCTTGTAAAGATGATCTTTGGAGATTATGCATAAAGAGGCTAAGGACACCAAAACGGC 456  
Db |||||  
QY 762 TCGGCTTGTAAAGTCTATTCTGGAGATTATGCATAAAGAGCTCAAGATATAAAATTCAC 821  
Db |||||  
QY 457 TGACGAGGTTCCCTGAGATCCTGGCCCATATAACTTTGTAGGGGCTCTCATTTGGCAA 516  
Db |||||  
QY 822 AGAAGAGATCCCTTGAAGATTTTAGCTCATATAACTTTGTGGAGCTCTTATTTGGTAA 881  
Db |||||  
QY 517 GGAAGGACGGAACCTGAAGAGGTAGAGCAAGATACCGAGACAAAAATCACCATCTCCTC 576  
Db |||||  
QY 882 AGAAGGAAGAATCTTAATAAAATTTAGCAAGACACACATAAATCACGATATCTCC 941  
Db |||||  
QY 577 GTTGAAGACCTTACCCTTTTAAACCTTGAGAGGACCATCACTGTGAGGGGGCCATCGA 636  
Db |||||  
QY 942 ATTGCAGGAATTGACGCTGTATAATCCAGAACGCATTTACAGTTTAAAGGCAATGTGA 1001  
Db |||||  
QY 637 GAATTGTTGACGGCCGAGGAGGAATAATCAAGAAAGTTTCGGGAGGCTTATGAGATGA 696  
Db |||||  
QY 1002 GACATGTGCCAAAGCTGAGGAGGAGATCTGAGAGAAATCAGGGAGTCTTTATGAAATGA 1061  
Db |||||  
QY 697 TGTGGCTGCCATGA-----GCTCTCACTGATCCCTGGGCTGAACCTGCTGTAGG 750  
Db |||||  
QY 1062 TATTGCTTCTATGAATCTTCAAGCACATTTAAATCTTGGATTAAATCTGAACGCTTGGG 1121  
Db |||||  
QY 751 TCTTTTCCAGCTTCATCCAGCGCAGTCCCGCCCTCCAGCAGCGTTACTGSGGCTGC 810  
Db |||||  
QY 1122 TCTGTTCACCCACTTCAGGAGTGCCACTCCCACTCAGGGCCCCCTTCAGCCATGAC 1181  
Db |||||  
QY 811 TCCCTATAGCTCCTTTATGCAAGCTCCCGAGCAGAGATGGTGCAGGTGTTTATCCCGC 870  
Db |||||  
QY 1182 TCCTCCCTACCGCAGTTTGAGCAATCAGAAACGGAGACTGTTCACTGTTTATCCCGC 1241  
Db |||||  
QY 871 CCAGGAGTGGCGGCATCATCGCAAGAGGGGCGACACATCAACAGCTCTCCCGTT 930  
Db |||||  
QY 1242 TCTATCAGTGGTGGCATCATCGCAAGAGGGGCGACACATCAAGCAGCTTTCTCGCT 1301  
Db |||||  
QY 931 TGCAGCGCTCCATCAAGATTGCAACCCGAAACACCTGACTCCAAAGTTGCTATGTT 990  
Db |||||  
QY 1302 TGCTGGAGCTTCAATTAAGATTGCTCCAGCGAGACACAGATGCTTAAAGTGAGGATGTT 1361  
Db |||||  
QY 991 TATCATCTAGTGGACCGCCAGAGCCCAATTCAAGGCTCAGGGAAGATCTATGGCAACT 1050  
Db |||||  
QY 1362 GATTATCACTGGACCAACAGAGGCTCAGTTCAAGGCTCAGGGAAGAAATTTATGGAATAAT 1421  
Db |||||  
QY 1051 CAAGGAGAGAACTTCTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACCCACATACGTT 1110  
Db |||||  
QY 1422 TAAAGAGAGAAACTTTGTTAGTCTTAAGAGAGGTGAACCTTGAAGCTCATATCAGAT 1481  
Db |||||  
QY 1111 GCCAGCATCAGCAGCTGCGGGTCAATTTGGCAAGGTGGAATAACCGTGAACGAGTTGCA 1170  
Db |||||  
QY 1482 GCCATCCTTGTCTGTCAGAGTTATTGGAAGAGGAGGCAAAACCGTGAATGAATTTCA 1541  
Db |||||  
QY 1171 GAATTTGAGCGCAGCTGAGGTGTTAGTACCAAGAGACGACCCCTGATGAGAACGACCA 1230  
Db |||||  
QY 1542 GAATTTGTCAAGTGCAGAGTTGTTGCTTCTGAGCAGACACCTGATGAGAAATGACCA 1601  
Db |||||  
QY 1231 GGTCACTCGTAAATCATCGGACATTTCTATGCGAGTCAATGGCTCAACGGAAGATCCG 1290  
Db |||||  
QY 1602 AGTGGTGTCAAAATAACTGCTACTTCTATGCTTGCAGGTTGCCAGAGGAAATTTCA 1661  
Db |||||  
QY 1291 AGACATCTGCGCCAGGTTAGCAGCAGCATCAAG 1327  
Db |||||  
QY 1662 GGAATTTCTGACTCAGGTAAGAGCAGCACCACACAG 1698  
Db |||||

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 13, 2004, 14:50:27 ; Search time 699 Seconds  
(without alignments)  
10380.433 Million cell updates/sec

Title: US-09-270-437D-5  
Perfect score: 1708  
Sequence: 1 agggagcgtgcgcacggcc.....atttccttcaggttttaaaa 1708

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_23Jan04: \*  
1: Geneseqn1980s: \*  
2: Geneseqn1990s: \*  
3: Geneseqn2000s: \*  
4: Geneseqn2001as: \*  
5: Geneseqn2001bs: \*  
6: Geneseqn2002s: \*  
7: Geneseqn2003as: \*  
8: Geneseqn2003bs: \*  
9: Geneseqn2003cs: \*  
10: Geneseqn2004s: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description        |
|------------|--------|-------------|--------|-------|--------------------|
| 1          | 1703   | 99.7        | 1708   | 3     | Aaz36151 DNA encod |
| 2          | 1669.2 | 97.7        | 1946   | 3     | Aaz36153 An altern |
| 3          | 1389   | 81.3        | 2780   | 4     | Aak94782 Human ful |
| 4          | 1161.8 | 68.0        | 2224   | 2     | Aaz10617 cDNA enco |
| 5          | 1085.4 | 63.5        | 1182   | 4     | Aas26148 Human cDN |
| 6          | 1085.4 | 63.5        | 1182   | 7     | Abx73489 Human nov |
| 7          | 1069.8 | 62.6        | 1129   | 4     | Aas26566 Human cDN |
| 8          | 1069.8 | 62.6        | 1129   | 7     | Abx73907 Human nov |
| 9          | 732.8  | 42.9        | 833    | 4     | Aak91969 Human cDN |
| 10         | 732.8  | 42.9        | 833    | 4     | Aak93655 Human cDN |
| 11         | 582.2  | 34.1        | 1740   | 3     | Aac66035 Human lun |
| 12         | 582.2  | 34.1        | 1740   | 6     | Abi49254 Human lun |
| 13         | 582.2  | 34.1        | 1740   | 6     | Abi49254 Human lun |
| 14         | 582.2  | 34.1        | 1743   | 6     | Abi49299 Human lun |
| 15         | 582.2  | 34.1        | 1743   | 6     | Abi49297 Human lun |
| 16         | 582.2  | 34.1        | 1743   | 6     | Abg92483 Human lun |
| 17         | 582.2  | 34.1        | 1743   | 6     | Abg92485 Human lun |
| 18         | 582.2  | 34.1        | 1743   | 8     | Ada28540 Recombina |
| 19         | 582.2  | 34.1        | 1743   | 8     | Ada28537 Recombina |
| 20         | 582.2  | 34.1        | 1764   | 6     | Abi49283 Human lun |
| 21         | 582.2  | 34.1        | 1764   | 6     | Abg92469 Human lun |
| 22         | 582.2  | 34.1        | 1764   | 8     | Ada28518 Recombina |
| 23         | 580.6  | 34.0        | 1740   | 8     | Ada28437 Human lun |

|    |       |      |      |   |                    |
|----|-------|------|------|---|--------------------|
| 24 | 580.6 | 34.0 | 4159 | 3 | Aaz36150 DNA encod |
| 25 | 580.6 | 34.0 | 4181 | 3 | Aac65900 Human lun |
| 26 | 580.6 | 34.0 | 4181 | 6 | Abi65405 Lung canc |
| 27 | 580.6 | 34.0 | 4181 | 6 | Abi49119 Human lun |
| 28 | 580.6 | 34.0 | 4181 | 6 | Abg92305 Human lun |
| 29 | 580.6 | 34.0 | 4181 | 8 | Ada28265 Human lun |
| 30 | 580.6 | 34.0 | 4181 | 9 | Adi4661 Human src  |
| 31 | 580.6 | 34.0 | 4181 | 9 | Ades3470 Human lun |
| 32 | 580.6 | 34.0 | 4434 | 8 | Acho3940 Human cDN |
| 33 | 580.6 | 34.0 | 4601 | 6 | AbA99958 Human KOC |
| 34 | 551.4 | 32.3 | 1985 | 5 | Aas76779 DNA encod |
| 35 | 530   | 31.0 | 4264 | 5 | Aas86150 DNA encod |
| 36 | 504   | 29.5 | 3283 | 3 | Aaz36154 An altern |
| 37 | 504   | 29.5 | 3667 | 5 | Aas70982 DNA encod |
| 38 | 489.2 | 28.6 | 2010 | 5 | AbS76442 cDNA enco |
| 39 | 489.2 | 28.6 | 3412 | 3 | Aaz36152 DNA encod |
| 40 | 489.2 | 28.6 | 3694 | 4 | Aas26150 Human cDN |
| 41 | 489.2 | 28.6 | 3694 | 7 | Abx73491 Human nov |
| 42 | 489.2 | 28.6 | 3694 | 7 | Abx73491 Human nov |
| 43 | 454   | 26.6 | 1707 | 7 | ACA90176 cDNA enco |
| 44 | 267.2 | 15.6 | 2290 | 4 | Aas26153 Human cDN |
| 45 | 267.2 | 15.6 | 2290 | 7 | Abx73494 Human nov |

ALIGNMENTS

RESULT 1  
AAZ36151  
ID AAZ36151 standard; DNA; 1708 BP.  
XX  
AC AAZ36151;  
XX  
DT 11-FEB-2000 (first entry)  
XX  
DE DNA encoding cancer associated antigen KOC-2.  
XX  
KW Cancer associated antigen; KOC-2; cancer; vaccine; CT7; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO9954738-A1.  
XX  
PD 28-OCT-1999.  
XX  
PF 16-MAR-1999; 99WO-US005766.  
XX  
PR 17-APR-1998; 98US-00061709.  
XX  
PA (LUDW-) LUDWIG INST CANCER RES.  
XX  
PI Chen Y, Gure A, Tsang S, Stockert E, Jager E, Knuth A, Old LJ;  
XX  
DR WPI; 2000-013284/01.  
XX  
PT Nucleotides representing cancer-associated genes, used to develop  
XX products for the diagnosis, monitoring and treatment of cancers.  
XX  
PS Claim 55; Page 40; 4pp; English.  
XX  
CC The present sequence represents a cancer associated antigen gene  
XX designated KOC-2. The specification also describes a cancer associated  
XX antigen designated CT7. The CT7 polynucleotide was isolated from SK-MEL-  
XX 37 melanoma cells. The polypeptide has some homology with MAGE-10,  
XX limited to about 210 carboxy terminal amino acids. The amino terminal of  
XX the protein has a repetitive pattern, with repeats rich in serine,  
XX proline, glutamine and leucine, and an almost invariable core of the  
XX peptide given in AAZ43877. The CT7 polypeptide can be processed to  
XX peptides which provoke lysis by cytolytic T cells. The polynucleotides  
XX and polypeptides can be used for treating a cancerous condition and  
XX screening for or diagnosing cancerous conditions. The cancer associated  
XX antigens can be used as an immunogenic or vaccine composition with an  
XX adjuvant, e.g. a cytokine, a saponin, or granulocyte macrophage-colony

CC stimulating factor (GM-CSF)  
 XX Sequence 1708 BP; 447 A; 469 C; 473 G; 314 T; 0 U; 5 Other;  
 SQ

Query Match 99.7%; Score 1703; DB 3; Length 1708;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGACGCTGCGCAGCCCGCCAGTTTACCCGGGGGCCATCATGAGCTGAATGGCCA 60  
 DB 1 AGGGACGCTGCGCAGCCCGCCAGTTTACCCGGGGGCCATCATGAGCTGAATGGCCA 60

QY 61 CCAGTTGGAGAACCATGCGCTGAAGGTCTCTACATCCCGGATGAGCAGATAGACAGGG 120  
 DB 61 CCAGTTGGAGAACCATGCGCTGAAGGTCTCTACATCCCGGATGAGCAGATAGACAGGG 120

QY 121 ACCTGAGATGGGCGCGAGGGGCTTTGGGTCTCGGGGTGAGCCCGCAGGGCTCAC 180  
 DB 121 ACCTGAGATGGGCGCGAGGGGCTTTGGGTCTCGGGGTGAGCCCGCAGGGCTCAC 180

QY 181 TGTGTCAGCGGGGCGCCAGCAAGCAGCAGCAAGTGGACATCCCGCTTCGGCTCCTGGT 240  
 DB 181 TGTGTCAGCGGGGCGCCAGCAAGCAGCAGCAAGTGGACATCCCGCTTCGGCTCCTGGT 240

QY 241 GCCCACCCAGTATGTGGTGCCATTTATTTGGCAAGGGGGCCACCATCCGCAACATCAC 300  
 DB 241 GCCCACCCAGTATGTGGTGCCATTTATTTGGCAAGGGGGCCACCATCCGCAACATCAC 300

QY 301 AAAACAGACCCAGTCCAGATAGAGTGCATAGGAAGGAGAGCAGGTGCAGCTGAAA 360  
 DB 301 AAAACAGACCCAGTCCAGATAGAGTGCATAGGAAGGAGAGCAGGTGCAGCTGAAA 360

QY 361 AGCCATCAGTGTGCACTCCACCCCTGAGGGTGTCTCCCTCCGCTTGAAGATGATCTTGA 420  
 DB 361 AGCCATCAGTGTGCACTCCACCCCTGAGGGTGTCTCCCTCCGCTTGAAGATGATCTTGA 420

QY 421 GATTATGATTAAGAGGTAGGACCAACACCGCTGACGAGTTCCCTTGAAGATCCT 480  
 DB 421 GATTATGATTAAGAGGTAGGACCAACACCGCTGACGAGTTCCCTTGAAGATCCT 480

QY 481 GGCCCAATAACTTTGTAGGCGCTCTCATTTGGCAAGGAGGACGGAACCTTGAAGAAGGT 540  
 DB 481 GGCCCAATAACTTTGTAGGCGCTCTCATTTGGCAAGGAGGACGGAACCTTGAAGAAGGT 540

QY 541 AGAGCAAGATACCGAGACAAATACCATCTCTCGTTTGAAGACCTTACCTTTACAA 600  
 DB 541 AGAGCAAGATACCGAGACAAATACCATCTCTCGTTTGAAGACCTTACCTTTACAA 600

QY 601 CCTGAGAGGACCATCACTGTGAAGGGGCCATCGAGAAATGTTGCGGGCGGAGCAGGA 660  
 DB 601 CCTGAGAGGACCATCACTGTGAAGGGGCCATCGAGAAATGTTGCGGGCGGAGCAGGA 660

QY 661 AATAATGAAGAAAGTTGGAGGCGCTATGAGATGATGTGGCTGCATGAGCTCTCACCT 720  
 DB 661 AATAATGAAGAAAGTTGGAGGCGCTATGAGATGATGTGGCTGCATGAGCTCTCACCT 720

QY 721 GATCCCTGGCTGAACCTGGCTGTAGTGTCTTTTCCAGCTTCATCCGCGCAGTCCC 780  
 DB 721 GATCCCTGGCTGAACCTGGCTGTAGTGTCTTTTCCAGCTTCATCCGCGCAGTCCC 780

QY 781 GCCGCTCCAGCAGCGTTACTGGGGCTGCTCCCTATAGTCCCTTTATGCAAGGCTCCGA 840  
 DB 781 GCCGCTCCAGCAGCGTTACTGGGGCTGCTCCCTATAGTCCCTTTATGCAAGGCTCCGA 840

QY 841 GCAGAGATGTCAGAGTGTATATCCCGCCAGCAGTGGCGCCATCATCGGCAGAA 900  
 DB 841 GCAGAGATGTCAGAGTGTATATCCCGCCAGCAGTGGCGCCATCATCGGCAGAA 900

QY 901 GGGGCGACATCAACAGCTCTCCCGGTTTGGCAGCGCTCCCATCAAGATTGACCAACC 960  
 DB 901 GGGGCGACATCAACAGCTCTCCCGGTTTGGCAGCGCTCCCATCAAGATTGACCAACC 960

QY 961 CGAAACCTGACTCAAGTTCGTATGTTATCATCTGACCGCCGAGGGCCCAATT 1020

DB 961 CGAAACCTGACTCCAAAGTTTCGTATGTTATCATCTGACCCGAGAGGCCCAATT 1020

QY 1021 CAAGGCTCAGGAGAGAAATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGGA 1080  
 DB 1021 CAAGGCTCAGGAGAGAAATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGGA 1080

QY 1081 GGAAGTGAAGCTGGAGACCCACATACATCGTGTGCCAGCATCAGCAGCTGSCCGGTCATTGG 1140  
 DB 1081 GGAAGTGAAGCTGGAGACCCACATACATCGTGTGCCAGCATCAGCAGCTGSCCGGTCATTGG 1140

QY 1141 CAAAGTGGAAGAAACGGTGAACAGTTGAGCAATTTGACGGCAGCTGAGGTGGTAGTACC 1200  
 DB 1141 CAAAGTGGAAGAAACGGTGAACAGTTGAGCAATTTGACGGCAGCTGAGGTGGTAGTACC 1200

QY 1201 AAGAGACCAAGACCCCTGATGAGAACACAGGTCTCATCGTGAATAATCATCGGACATTTCTA 1260  
 DB 1201 AAGAGACCAAGACCCCTGATGAGAACACAGGTCTCATCGTGAATAATCATCGGACATTTCTA 1260

QY 1261 TGCAGTCTAGATGGCTCAACGGAAGATCCGACATCTCTGGCCAGGTTAAGCAGCAGCA 1320  
 DB 1261 TGCAGTCTAGATGGCTCAACGGAAGATCCGACATCTCTGGCCAGGTTAAGCAGCAGCA 1320

QY 1321 TCAGAGGGACAGAGTTAAACAGGCCAGCAGCGAGGAGTGCACAGCCCTCCCTGTGCC 1380  
 DB 1321 TCAGAGGGACAGAGTTAAACAGGCCAGCAGCGAGGAGTGCACAGCCCTCCCTGTGCC 1380

QY 1381 CTTCAGTCCAGGACAAACACCGGAGAAATCGAGAGTGTGCTCTCCCGCGCAGGCTGA 1440  
 DB 1381 CTTCAGTCCAGGACAAACACCGGAGAAATCGAGAGTGTGCTCTCCCGCGCAGGCTGA 1440

QY 1441 GAATGATCGGAATCCGGGACACNTGGCCGGGCTGTAGATCAGTGTTCGCCCATGATT 1500  
 DB 1441 GAATGATCGGAATCCGGGACACNTGGCCGGGCTGTAGATCAGTGTTCGCCCATGATT 1500

QY 1501 GAGAAAGATGTTCCAGTGAAGAACCTGATCTNTCAGCCCAACCAACCCCAATTTGCC 1560  
 DB 1501 GAGAAAGATGTTCCAGTGAAGAACCTGATCTNTCAGCCCAACCAACCCCAATTTGCC 1560

QY 1561 CCAACACTGTTNTGCCCTCGGGGTGTGAGAAATTTAGCGCAAGGACATTTTAAACGTTGG 1620  
 DB 1561 CCAACACTGTTNTGCCCTCGGGGTGTGAGAAATTTAGCGCAAGGACATTTTAAACGTTGG 1620

QY 1621 ATTGTTTAAAGAGCTCTCCAGGCCCAACCAAGAGGTTGATCACAACCTCAGTGGGAAGA 1680  
 DB 1621 ATTGTTTAAAGAGCTCTCCAGGCCCAACCAAGAGGTTGATCACAACCTCAGTGGGAAGA 1680

QY 1681 AAAATAAAATTTCTTCAGGTTTAAAA 1708  
 DB 1681 AAAATAAAATTTCTTCAGGTTTAAAA 1708

## RESULT 2

AAZ36153  
 ID AAZ36153 standard; DNA; 1946 BP.

XX AC AAZ36153;

XX AC AAZ36153;

DT 11-FEB-2000 (first entry)

XX An alternative form of DNA encoding cancer associated antigen KOC-2.

DE Cancer associated antigen; KOC-2; cancer; vaccine; CT7; ss.

XX Homo sapiens.

XX WO9954738-A1.

XX 28-OCT-1999.

XX 16-MAR-1999; 99WO-US005766.

XX 17-APR-1998; 98US-00061709.

XX (LUDW-) LUDWIG INST CANCER RES.  
XX Chen Y, Gure A, Tsang S, Stockert E, Jager E, Knuth A, Old LJ;  
XX WPI; 2000-013284/01.  
XX Nucleotides representing cancer-associated genes, used to develop  
XX products for the diagnosis, monitoring and treatment of cancers.  
XX Claim 55; Page 42; 44pp; English.  
XX The present sequence represents an alternative form of a cancer  
XX associated antigen gene designated KOC-2. The specification also  
XX describes a cancer associated antigen designated CT7. The CT7  
XX polynucleotide was isolated from SK-MEL-37 melanoma cells. The  
XX polypeptide has some homology with MAGS-10, limited to about 210 carboxy  
XX terminal amino acids. The amino terminal of the protein has a repetitive  
XX pattern, with repeats rich in serine, proline, glutamine and leucine, and  
XX an almost invariable core of the peptide given in AAY43877. The CT7  
XX polypeptide can be processed to peptides which provoke lysis by cytolytic  
XX T cells. The polynucleotides and polypeptides can be used for treating a  
XX cancerous condition and screening for or diagnosing cancerous conditions.  
XX The cancer associated antigens can be used as an immunogenic or vaccine  
XX composition with an adjuvant, e.g. a cytokine, a saponin, or granulocyte  
XX macrophage-colony stimulating factor (GM-CSF)  
XX SQ Sequence 1946 BP; 502 A; 528 C; 553 G; 358 T; 0 U; 5 Other;

Query Match 97.7%; Score 1669.2; DB 3; Length 1946;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1676; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 30 CCGCGGGAGCCATCATGAAGCTGAATGGCCACAGCTGGAGAACCATGCCCTGAAGGTCT 89  
DB 268 CCAGGCAAGCCATCATGAAGCTGAATGGCCACAGCTGGAGAACCATGCCCTGAAGGTCT 327

QY 90 CTTACATCCCGATGAGCAGATAGCACAGGGACCTGAGAACTGGGCGCCGAGGGGCTTTG 149  
DB 328 CTTACATCCCGATGAGCAGATAGCACAGGGACCTGAGAACTGGGCGCCGAGGGGCTTTG 387

QY 150 GCTCTCGGGTCAGCCCGCCAGGGCTCACTGTGGCAGCGGGGCCCCAGCCAGCAGC 209  
DB 388 GCTCTCGGGTCAGCCCGCCAGGGCTCACTGTGGCAGCGGGGCCCCAGCCAGCAGC 447

QY 210 AGCAAGTGACATCCCTCTGGCTCCTGGTGGCCACCCAGTATGGGTGCCATTAATG 269  
DB 448 AGCAAGTGACATCCCTCTGGCTCCTGGTGGCCACCCAGTATGGGTGCCATTAATG 507

QY 270 GCAAGAGGGGCGCCACCATCCGCAATCAACAAACAGACCCAGTCCAGATAGACGTGC 329  
DB 508 GCAAGAGGGGCGCCACCATCCGCAATCAACAAACAGACCCAGTCCAGATAGACGTGC 567

QY 330 ATAGGAAGGAGAACGACGTGAGCTGAAAGAGCCATCAGTGTGCACTCCACCCCTGAGG 389  
DB 568 ATAGGAAGGAGAACGACGTGAGCTGAAAGAGCCATCAGTGTGCACTCCACCCCTGAGG 627

QY 390 GCTGCTCTCCGCTTGAAGTATCTTGGAGATATGCAATAAGAGGCTAAGGACACCA 449  
DB 628 GCTGCTCTCCGCTTGAAGTATCTTGGAGATATGCAATAAGAGGCTAAGGACACCA 687

QY 450 AAACGGCTGACGAGGTCCCTGAGAGTCCCTGGCCCAATAAATTTGTAGGGGTCTCA 509  
DB 688 AAACGGCTGACGAGGTCCCTGAGAGTCCCTGGCCCAATAAATTTGTAGGGGTCTCA 747

QY 510 TTGGCAAGGAGGAGCGAACTTGAAAGAGGTAGAGCAAGATACCGAGACAAAATCAACA 569  
DB 748 TTGGCAAGGAGGAGCGAACTTGAAAGAGGTAGAGCAAGATACCGAGACAAAATCAACA 807

QY 570 TCTCTCTGTCGAGACCTTACCTTTACACCTTGAGAGACCATCACTGTGAAGGGG 629  
DB 808 TCTCTCTGTCGAGACCTTACCTTTACACCTTGAGAGACCATCACTGTGAAGGGG 867

QY 530 CCATCGAGAAATTTGTCAGGGCCGAGCAGGAAATTAATGAAGAAAGTTCCGGAGGCCTATG 689  
DB 868 CCATCGAGAAATTTGTCAGGGCCGAGCAGGAAATTAATGAAGAAAGTTCCGGAGGCCTATG 927

QY 690 AGAATGATGTGGTGCCATGAGCTCTCACCTGATCCCTGGCCTGAACCTGGCTGCTGTAG 749  
DB 928 AGAATGATGTGGTGCCATGAGCTCTCACCTGATCCCTGGCCTGAACCTGGCTGCTGTAG 987

QY 750 GTCTTTTCCAGAGTTCAATCCAGGCGAGTCCCGCGCTCCACAGAGGTTACTTGGGGCTG 809  
DB 988 GTCTTTTCCAGAGTTCAATCCAGGCGAGTCCCGCGCTCCACAGAGGTTACTTGGGGCTG 1047

QY 810 CTCCTATAGCTCTCTTTATGACGCTCCCGAGCAGAGATGGTGCAGGTGTTATCCCGC 869  
DB 1048 CTCCTATAGCTCTCTTTATGACGCTCCCGAGCAGAGATGGTGCAGGTGTTATCCCGC 1107

QY 870 CCCAGGAGTGGGCGCATCATCGGAAGAGGGGAGCAGACATCAACAGCTCTCCCGGT 929  
DB 1108 CCCAGGAGTGGGCGCATCATCGGAAGAGGGGAGCAGACATCAACAGCTCTCCCGGT 1167

QY 930 TTCCGAGCGCTCCATCAGATTGCAACCCGACCCGACACCTGACTCCAAAGTTGGTATGG 989  
DB 1168 TTCCGAGCGCTCCATCAGATTGCAACCCGACCCGACACCTGACTCCAAAGTTGGTATGG 1227

QY 990 TTATCATCACTGACCGCCAGAGGCCCAATTCAGGCTCAGGGAAGAATCTATGGCAAC 1049  
DB 1228 TTATCATCACTGACCGCCAGAGGCCCAATTCAGGCTCAGGGAAGAATCTATGGCAAC 1287

QY 1050 TCAAGGAGGAGAACTTTTGTGTCCTCAGAGGAGGAGTGAAGCTGGAGACCCCATAGCTG 1109  
DB 1288 TCAAGGAGGAGAACTTTTGTGTCCTCAGAGGAGGAGTGAAGCTGGAGACCCCATAGCTG 1347

QY 1110 TGCAGCATCAGCAGCTGGCGGCTCATTTGGCAAGAGTGGAAAAACGTTCAACGAGTTGC 1169  
DB 1348 TGCAGCATCAGCAGCTGGCGGCTCATTTGGCAAGAGTGGAAAAACGTTCAACGAGTTGC 1407

QY 1170 AGAATTGACGGCAGCTGAGTGTAGTACCAAGAGACCAAGACCCCTGTATGAGAACGACC 1229  
DB 1408 AGAATTGACGGCAGCTGAGTGTAGTACCAAGAGACCAAGACCCCTGTATGAGAACGACC 1467

QY 1230 AGGTCTCGTGAATATCATCGGACATTTCTATCCAGTCAAGTGGCTCAACGGAAGTCC 1289  
DB 1468 AGGTCTCGTGAATATCATCGGACATTTCTATCCAGTCAAGTGGCTCAACGGAAGTCC 1527

QY 1290 GAGACATCTCTGGCCCGAGTTAAGCAGCAGCATCAGAAAGGAGCAGAGTAACACGCGCCAGG 1349  
DB 1528 GAGACATCTCTGGCCCGAGTTAAGCAGCAGCATCAGAAAGGAGCAGAGTAACACGCGCCAGG 1587

QY 1350 CAGGAGGAGTGAACAGACCCCTCTCTGTCCTTNGAGTCCAGGACAAACACGGGAGAA 1409  
DB 1588 CAGGAGGAGTGAACAGACCCCTCTCTGTCCTTNGAGTCCAGGACAAACACGGGAGAA 1647

QY 1410 ATCGAGAGTGTCTCTCCCGCAGCGCTGAGAAATGAGTGGGAATCCGGGACACNTGGGC 1469  
DB 1648 ATCGAGAGTGTCTCTCCCGCAGCGCTGAGAAATGAGTGGGAATCCGGGACACNTGGGC 1707

QY 1470 CGGCTGTAGATCAGGTTTGCCCATTTGATTTGAGAAAGATGTTCCAGTGAAGAACCTGA 1529  
DB 1708 CGGCTGTAGATCAGGTTTGCCCATTTGATTTGAGAAAGATGTTCCAGTGAAGAACCTGA 1767

QY 1530 TCTNTCAGCCCAACACACCCCAATTCGCCCAACACTGTTNGCCCTCGGGGTGTCAG 1589  
DB 1768 TCTNTCAGCCCAACACACCCCAATTCGCCCAACACTGTTNGCCCTCGGGGTGTCAG 1827

QY 1590 AAATTTAGCGCAAGGACCTTTTAAACGTGGATTGTTTAAAGAGGCTCTCCAGGCCCCAC 1649  
DB 1828 AAATTTAGCGCAAGGACCTTTTAAACGTGGATTGTTTAAAGAGGCTCTCCAGGCCCCAC 1887

QY 1650 CAGAGGGTGGATCAGACCTCAGTGGGAGAGAAAAATAAATTTCTTCAAGTTTAAAA 1708  
DB 1888 CAGAGGGTGGATCAGACCTCAGTGGGAGAGAAAAATAAATTTCTTCAAGTTTAAAA 1946

```

RESULT 3
AAK94782
ID AAK94782 standard; cDNA; 2780 BP.
XX
AC AAK94782;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human full-length cDNA; cDNA synthesis; oligo-capping; ss.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS Homo sapiens.
XX
PN EP1130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-00114089.
XX
PR 08-JUL-1999; 99JP-00194486.
XX
PR 11-JAN-2000; 2000JP-00118774.
XX
PR 02-MAY-2000; 2000JP-00183765.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI: 2001-524255/58.
XX
DR P-PSDB; AAW93826.
XX
XX
830 Primers useful for synthesizing full length cDNA clones and their use
in genetic manipulation.
XX
Claim 8; SEQ ID NO 3886; 1380pp + Sequence Listing; English.
XX
XX
The invention relates to primers for synthesizing full length cDNA
clones. 830 cDNA molecules encoding a human protein have been isolated
and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
been determined. Primers for synthesizing the full length cDNA are useful
for clarifying the function of the protein encoded by the cDNA. The full
length clones were obtained by construction of full length enriched cDNA
libraries that were synthesised by the oligo-capping method. The primers
enable the production of the full length cDNA easily without any special
methods. The present sequence is a full length human cDNA of the
invention. Note: The sequence data for this patent did not form part of
the printed specification, but was obtained in CD-ROM format directly
from EPO
XX
SQ Sequence 2780 BP; 768 A; 681 C; 671 G; 560 T; 0 U; 0 Other;
XX
Query Match 81.3%; Score 1389; DB 4; Length 2780;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1405; Conservative 0; Mismatches 5; Indels 6; Gaps 1;
XX
299 ACAAAACAGACCCAGCTCCAGATAGACGTGCATAGGAGGAGAAACGCGAGTGCAGCTGAA 358
1 ACAAAACAGACCCAGCTCCAGATAGACGTGCATAGGAGGAGAAACGCGAGTGCAGCTGAA 60
359 AAAGCCATCAGTGTGCACTCCACCCCTGAGGGCTGCTCCTCCGCTGTGAAGATGATCTTG 418
61 AAAGCCATCAGTGTGCACTCCACCCCTGAGGGCTGCTCCTCCGCTGTGAAGATGATCTTG 120
419 GAGATTATGATTAAGAGCTTAAGACACCAAAACGGGTGAGAGTTCCTCCGAAGATC 478
121 GAGATTATGATTAAGAGCTTAAGACACCAAAACGGGTGAGAGTTCCTCCGAAGATC 180
479 CTGGCCCATATAACTTTGTAGGGCGTCTCATTGGCAAGGAGGACCGAACTTGAAGAAG 538
181 CTGGCCCATATAACTTTGTAGGGCGTCTCATTGGCAAGGAGGACCGAACTTGAAGAAG 240
539 GTAGAGCAGATACCGAGACAAAATACCATCTCCTGTTGCAAGACCTTACCCCTTAC 598

```

Db 1321 AACGTGGATTGTTTAAAGAGCTCTCCAGGCCCCACCAAGAGGGTGGATCACACCTCAG 1380  
Qy 1673 TGGGAAGAAAATAAAATTTCTTCAGGTTTAAAA 1708  
Db 1381 TGGGAAGAAAATAAAATTTCTTCAGGTTTAAAA 1416

RESULT 4  
AAZ10617  
ID AAZ10617 standard; cDNA; 2224 BP.  
XX AC AAZ10617;  
XX DT 17-NOV-1999 (first entry)  
XX DE cDNA encoding a murine c-myc coding region determinant binding protein.  
XX KW c-myc coding region determinant binding protein; CRD-BP; tumor; c-myc;  
XX KW endonucleolytic attack; half-life; breast cancer; colon cancer;  
XX KW pancreatic cancer; ss.  
XX OS Mus musculus.  
XX FH Key Location/Qualifiers  
XX CDS 131..1864  
XX FT /\*tag= a  
XX PN WO9946594-A2.  
XX PD 16-SEP-1999.  
XX PP 05-MAR-1999; 99WO-US004897.  
XX PR 09-MAR-1998; 98US-0077372P.  
XX PA (WISC ) WISCONSIN ALUMNI RES FOUND.  
XX PI Ross J;  
XX DR WPI; 1999-551506/46.  
XX DR P-PSDB; AAY30649.  
XX PT Diagnosing presence or absence of a tumor in a human by examining c-myc  
XX PT coding region determinant-binding protein.  
XX PS Example; Fig 1A-D; 79pp; English.  
XX SS The present sequence encodes a murine c-myc coding region determinant  
XX CC binding protein (CRD-BP). The presence or absence of a tumor can be  
XX CC determined by determining the levels of CRD-BP present in the suspect  
XX CC tissue, where the CRD-BP shields c-myc RNA from endonucleolytic attack  
XX CC and so prolongs its half-life. The methods are used for diagnosing  
XX CC presence or absence of a tumor in a human, especially breast, colon and  
XX CC pancreatic cancer. They are also used to inhibit cancer cell growth  
XX SQ Sequence 2224 BP; 580 A; 609 C; 618 G; 417 T; 0 U; 0 Other;

Query Match 68.0%; Score 1161.8; DB 2; Length 2224;  
Best Local Similarity 82.7%; Pred. No. 0;  
Matches 1409; Conservative 0; Mismatches 262; Indels 33; Gaps 6;  
Qy 30 CCCGGGAGCCATCATGAAGCTGAATGGCCACCACTTGGAGAACCATGCCCTGAAGTCT 89  
Db 525 CAGGCAAGCTATCATGAAGCTAAATGGCCATCACTGGAGAACCATGCCCTGAAGTCT 584  
Qy 90 CCTACATCCCCGATGAGCAGATAGCACAGGACCTGAGAAATGGCGCGGAGGGGGCTTG 149  
Db 585 CCTACATACCTGATGAGCAGATAACACAAGGTCCTGAGAAATGGCGCTGTGAGGCTTG 644  
Qy 150 GCTCTCGGGGTTCAGCCCGCAGGGCTCACTGTGGAGCGGGGGCCCCCAGCCAGCAGC 209  
Db 645 GGTCTCGGGGCGAGCCCGCAAGGGTCGCCCGTGGCAGAGGGGCTCCAGCCCAAGCAGC 704

Qy 210 AGCAATGTGACATCCCCCTTCGGCTCCTGGTCCACCCAGTATGTGGTGCATTATTG 269  
Db 705 AGCAGTGGACATCCCTCTCCGGCTCCTGGTCCACCCAGTATGTAGCGGTATCATTTG 764  
Qy 270 GCAAGGAGGGGCGACCATCCGCAACATCAAAAACAGACCAGTCCAAAGATAGAGTGC 329  
Db 765 GCAAGGAGGGTGCACCATCCGAAACATCAAAAACAGACGACGATCCAAAATAGACGTG 824  
Qy 330 ATAGGAGGAGAACCGAGGTGCAGTGAAGAGCCATCAGTGTGCACTCCACCCCTGAGG 389  
Db 825 ATAGGAGGAGAAATCGGGCGCTCGGAGAGGCCATCAGCGTGCANTTCAACCCCTGAAG 884  
Qy 390 GCTGCTCCTCCGCTTGTAAAGATGATCTTGGAGATTATGCATTAAGAGGCTTAAGGACCA 449  
Db 885 GCTGCTCCTCCGCTGCAAGATGATCTTGGAGATTATGCACAAAGAGGCAAGGACCA 944  
Qy 450 AAACGCTGACAGGTTCCCTGAGATCTCGGCCATTAATTAACCTTTGTAGCGGTCTCA 509  
Db 945 AAACGCGCATGAAGTTCCCTGAAATCTCGCTCATAACTTCTGCGGCGGACTCA 1004  
Qy 510 TTGGCAAGGAAGGACGGAACCTGMAAGAGCTAGAGCAAGATACCGAGACAAAAATCACCA 569  
Db 1005 TTGGCAAGGAAGGCGGAACCTGMAAGAGGTGGAGCAGACACAGAGCAAGATCACCA 1064  
Qy 570 TCTCTCTGTGCAAGACCTTACCTTTTCAACCTTGAGAGGACCATCATGTGAAGGGG 629  
Db 1065 TCTCATGCTCCAGGACCTCAGGCTCTATAACCTGAGAGGACCATCATGTGAAGGGG 1124  
Qy 630 CCATCGAGAAATGTTGACGGGCGGAGCAGGAATAATGAAGAAAGTTCGGAGGGCTATG 689  
Db 1125 CCATTGAGAACTGTTGACGGGCGGAGCAGGAGATCATGAAGAAAGTTCGAGAGGCTTACG 1184  
Qy 690 AGAATGATGTGGTCCCATGAGC-----TCTCACCTGATCCTCGCTGAACTGGCTG 743  
Db 1185 AGAAGCAGTGGCGGCATGAGCTTGACGTCCACCTCATCCTCGGCTTAACTGGCTG 1244  
Qy 744 CTGTAGGTCTTTTCCAGCTTCATCCAGCGAGTCCCGCGGCTCCACGAGGCTTACTG 803  
Db 1245 CTTGAGGTCTCTTCCAGCTTCATCCAGCGCTGTCCTCTCTCCAGCAGTGTCTACTG 1304  
Qy 804 GGGCTGCTCCTATAGTCTCTTTATGACGGCTCCCGAGCAGGAGATGGTGAAGTGTTTA 863  
Db 1305 GGGCTGCTCCTATAGTCTCTTCATGACGGCTCCGAGCAGGAGATGGTGAAGTGTTC 1364  
Qy 864 TCCCGCCGAGCAGTGGCGGCGCATATCGGAAGAGGGGAGCAGCATCAAAACAGTCT 923  
Db 1365 TCCCGCCGAGGCTGTGGGCGGCATCATTTGGCAAGAGGGCCAGCAGCATCAAAACACTCT 1424  
Qy 924 CCGGTTTCCAGCGCTCCATCAGATTGCACCCGAAACACCTGACTCCAAAGTTTC 983  
Db 1425 CCGTTCGCCAGCGCTCCATCAGATTGCTCCACAGAAACACCTGACTCCAAAGTTTC 1484  
Qy 984 GTATGTTTATCATCACTGGACCCGAGGCGCCAAATTCAGGCTCAGGGAAGAATCTATG 1043  
Db 1485 GAATGTTGCTCATCACTGACCCCGAGGCTCAGTTCAAGGCTCAGGGAAGAATTTATG 1544  
Qy 1044 GCAAACTCAAGGAGGAGAACTTTTGGTCCCAAGAGGAGTGAAGCTGGAGACCCACA 1103  
Db 1545 GCAAACTAAAGAGAGAGAAATTTTGGTCCCAAGGAGAGTAAAGCTAGAGACCCACA 1604  
Qy 1104 TAGCTGTGCAGCATCAGCAGCTGGCGGGTCAATTTGGCAAAAGTGAAGAAAGCTGAAAG 1163  
Db 1605 TAGGGTTCCGGCTTCCAGCAGCGCGCGGTCTATCGCAAGGCGCAAAACGGTGAATG 1664  
Qy 1164 AGTTGAGAAATTTGACGGCAGCTGAGGTGTAGTACCAAGAGACAGACCCCTGTATGAGA 1223  
Db 1665 AGCTGAGAACTTGAATGAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAG 1724  
Qy 1224 ACCACCAAGTTCATCGTGAAGAAATCATCGGACATTTTCTATCCAGTCAAGTGGCTCAACGGA 1283  
Db 1725 ACACCAAGTTCATCGTGAAGAAATCATCGGACATTTTCTATCCAGCAGTGGCTCAACGGA 1784  
Qy 1284 AGATCCGAGACATCTCTGGCCCGAGGTTTAAGCAGCAGCATCAGAAGGGGACAGTAAACAGG 1343

Db 1785 AGATCCGAGACATCTGGCTCAGTTAGACACAGCACCAGAGGGACAGACACTGG 1844  
QY 1344 CCAGGCGCGGAGGAAGTGA-CCAGCCCTCCCTGTCCTTNGAGTCAGGACCAACG 1402  
Db 1845 CCAGGCGCGGAGGAAGTGA-CCAGCCCTCCCTGTCCTTNGAGTCAGGACCAACG 1904  
QY 1403 GCAGAA-----ATCGAGAGTGTCTCCCGGAGGAGCTGAG 1441  
Db 1905 AGAACAAGACAGTGGAGGGGGGTGAGGGCCGGTGTGTTTCCAGAGGAGCTGAG 1964  
QY 1442 AATGAGTGGGAATCCGGACACNTGGGCGGGCTGTAGATCAGGTTTCCCACTTGATTG 1501  
Db 1965 AATGAGTGGGAATCAGGG-CATTTGGGCTGGCTGGAGATCAGGTTTGCACACTGTATTG 2023  
QY 1502 AGAAGATGTTCCAGTGGAGGACCTGATCTNTCAGCCCAACACCCCAATTTGGCC 1561  
Db 2024 AGAACAATGTTCCAGTGGAGGATCTGATCTCTCGCCCCAA--TTGAGCCAGCTGCCA 2081  
QY 1562 CAACACTGNTGCCCTCGGGGTGTCAGAAATNTAGCGCAAGGCACTTTTAAACGTGGA 1621  
Db 2082 CAGCCACCCCTTGAATATCACCATTGCAATCATAGCTTGGGTGCTTTTAAACGTGGA 2141  
QY 1622 TGTGTTAAAGAGCTCTCAGGCCCCCAGAGAGGGTGGATCAACCTCAGTGGGAAGAA 1681  
Db 2142 TGTGCT--TGAAGTTCTCCAGCTCCATGGAAGGATGGTCCAGTGGGGAAGAG 2199  
QY 1682 AATAAATTTCTTCAGGTTTAA 1705  
Db 2200 AATAAATTTCTTCAGGTTTAA 2223  
  
RESULT 5  
AAS26148  
ID AAS26148 standard; cDNA; 1182 BP.  
XX  
AC AAS26148;  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Human cDNA encoding a novel secreted protein, Seq ID 327.  
XX  
KW Human; immunosuppressive; antiarthritic; ss; antirheumatic; cytostatic;  
KW cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;  
KW antibacterial; virucide; fungicide; opthalmological; vulnerary;  
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;  
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;  
KW cerebral ischaemia; angiogenesis; nervous system disorder;  
KW Alzheimer's disease; infection; ocular disorder; corneal infection;  
KW wound healing; epithelial cell proliferation; skin ageing; food additive;  
KW preservative; antiproliferative.  
XX  
OS Homo sapiens.  
XX  
FN WO200155322-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001341.  
XX  
PF 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216860P.  
  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 26-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 14-AUG-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225211P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225475P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 23-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.



PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259679P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
FI P-PSDB; AUI16161.  
DR WPI; 2001-488783/53.  
XX P-PSDB; AUI16161.  
XX  
PT New nucleic acid molecules encoding 461 human secreted proteins for  
PT diagnosing, preventing, treating or ameliorating medical conditions and  
PT used as food additives or preservatives.  
XX  
PS Claim 1; SEQ ID NO 327; 980pp; English.  
XX  
CC The invention relates to isolated nucleic acid molecules and their  
CC encoded secreted proteins. The nucleic acids and proteins are used to  
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used  
CC in diagnosing a pathological condition or susceptibility to a  
CC pathological condition. Antibodies to the proteins can also be used in  
CC alleviating symptoms associated with the disorders and in diagnostic  
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbant assays  
CC (ELISA). Disorders which are diagnosed or treated include autoimmune

CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.  
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac  
CC arrest, cerebrovascular disorders e.g. cerebral ischemia, angogenesis,  
CC nervous system disorders e.g. Alzheimer's disease, infections caused by  
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,  
CC and many other disorders listed in the specification. The polypeptides  
CC can also be used to aid wound healing and epithelial cell proliferation,  
CC to prevent skin aging due to sunburn, to maintain organs before  
CC transplantation, for supporting cell culture of primary tissues, to  
CC regenerate tissues and in chemotaxis. The polypeptides can also be used  
CC as a food additive or preservative to increase or decrease storage  
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,  
CC minerals, cofactors and other nutritional components. The present  
CC sequence encodes a novel secreted protein of the invention. Note: The  
CC sequence data for this patent did not form part of the printed

|                       |      |   |                     |           |              |
|-----------------------|------|---|---------------------|-----------|--------------|
| Query Match           |      | 63.5%;  | Score 1085.4;       | DB 4;     | Length 1182; |
| Best Local Similarity |      | 98.9%;  | Pred. No. 3.9e-297; |           |              |
| Matches 1102;         |      | Conservative 0;   | Mismatches 6;       | Indels 6; | Gaps 1;      |
| QY                    | 601  | CCCTGAGAGGACCATCACTGTGAGAGGGGCCATCGAGAATTGTTGCAGGGCCGACGAGCA  | 660                 |           |              |
| DB                    | 18   | CGCTGAGAGGACCATCACTGTGAGAGGGGCCATCGAGAATTGTTGCAGGGCCGACGAGCA  | 77                  |           |              |
| QY                    | 661  | ATAATGAAGAAAGTTCCGGAGGCCCTATGAGAAATGATGGCTGCATGAGC-----TC     | 714                 |           |              |
| DB                    | 78   | ATAATGAAGAAAGTTCCGGAGGCCCTATGAGAAATGATGGCTGCATGAGC-----TC     | 137                 |           |              |
| QY                    | 715  | TCACCTGATCCCTGGCCTGAACCTGGCTGCTAGGTCTTTTCCAGGCTTCATCCAGCGC    | 774                 |           |              |
| DB                    | 138  | TCACCTGATCCCTGGCCTGAACCTGGCTGCTAGGTCTTTTCCAGGCTTCATCCAGCGC    | 197                 |           |              |
| QY                    | 775  | AGTCCCGCGCCTCCAGCAGCGTTACTGGGGCTCTCCCTATAGTCTCTTTATGAGGC      | 834                 |           |              |
| DB                    | 198  | AGTCCCGCGCCTCCAGCAGCGTTACTGGGGCTCTCCCTATAGTCTCTTTATGAGGC      | 257                 |           |              |
| QY                    | 835  | TCCCGAGCAGGAGATGGTGCAGTGTTATCCCGCCAGCAGTGGGGCCCATCATCGG       | 894                 |           |              |
| DB                    | 258  | TCCCGAGCAGGAGATGGTGCAGTGTTATCCCGCCAGCAGTGGGGCCCATCATCGG       | 317                 |           |              |
| QY                    | 895  | CAAGAAGGGGCGACACATCAAAACAGCTCTCCGGTTTGCAGCGCTCCATCAAGATTGC    | 954                 |           |              |
| DB                    | 318  | CAAGAAGGGGCGACACATCAAAACAGCTCTCCGGTTTGCAGCGCTCCATCAAGATTGC    | 377                 |           |              |
| QY                    | 955  | ACGACCCGAAACACCTGACTCCAAAGTTGTTATCATCACTGACCCGCGAGAGGC        | 1014                |           |              |
| DB                    | 378  | ACGACCCGAAACACCTGACTCCAAAGTTGTTATCATCACTGACCCGCGAGAGGC        | 437                 |           |              |
| QY                    | 1015 | CCAATTTCAAGGCTCAGGGGAAATCTATGCGCAAACTCAAGGAGGAGAACTTTTGGTCC   | 1074                |           |              |
| DB                    | 438  | CCAATTTCAAGGCTCAGGGGAAATCTATGCGCAAACTCAAGGAGGAGAACTTTTGGTCC   | 497                 |           |              |
| QY                    | 1075 | CAAGGAGGAAGTGAAGCTGGAGACCCACATAGTGTGTCAGCATCAGCAGTGGCGGGT     | 1134                |           |              |
| DB                    | 498  | CAAGGAGGAAGTGAAGCTGGAGACCCACATAGTGTGTCAGCATCAGCAGTGGCGGGT     | 557                 |           |              |
| QY                    | 1135 | CATTGGCAAGGTGGAAGAACGTTGAGTTCAGAAATTTGACGGCAGCTGAGTGCT        | 1194                |           |              |
| DB                    | 558  | CATTGGCAAGGTGGAAGAACGTTGAGTTCAGAAATTTGACGGCAGCTGAGTGCT        | 617                 |           |              |
| QY                    | 1195 | AGTACCAAGAGACCAACCCCTGATGAGAACGACCAAGGTTCATCGTCAAAATCATCGACA  | 1254                |           |              |
| DB                    | 618  | AGTACCAAGAGACCAACCCCTGATGAGAACGACCAAGGTTCATCGTCAAAATCATCGACA  | 677                 |           |              |
| QY                    | 1255 | TTTCTATGCCAGTCAGATGGCTCAACGGAGAGTCCGAGACATCTTGGCCAGGTTAAGCA   | 1314                |           |              |
| DB                    | 678  | TTTCTATGCCAGTCAGATGGCTCAACGGAGAGTCCGAGACATCTTGGCCAGGTTAAGCA   | 737                 |           |              |
| QY                    | 1315 | GCAGCATCAGAGGGGACAGATAACAGGCCCGACGAGGAGGAGTGCACGACCCCTCC      | 1374                |           |              |
| DB                    | 738  | GCAGCATCAGAGGGGACAGATAACAGGCCCGACGAGGAGGAGTGCACGACCCCTCC      | 797                 |           |              |
| QY                    | 1375 | CTGTCCCTTNGAGTCCAGGACCAACACAGGGGCAAAATCGAGAGTGTGCTCTCCCGGCGAG | 1434                |           |              |



|    |      |   |      |
|----|------|---|------|
| QY | 715  | TCACCTGATCCCTGGCGCTGAAACCTGGCTGCTGTAGGTCCTTTCCAGAGCTTCATCCAGCGC | 774  |
| Db | 138  | TCACCTGATCCCTGGCGCTGAAACCTGGCTGCTGTAGGTCCTTTCCAGAGCTTCATCCAGCGC | 197  |
| QY | 775  | AGTCCCGCGCGCTCCCGAGAGCGCTTACTGGGGTGCTCCCTATAGTCTCTTTATGAGGC     | 834  |
| Db | 198  | AGTCCCGCGCGCTCCCGAGAGCGCTTACTGGGGTGCTCCCTATAGTCTCTTTATGAGGC     | 257  |
| QY | 835  | TCCGAGCAGGAGATGGTGCAGGTGTTTATTCCTCCGCCAGGCGAGTGGGGCCCATCATCGG   | 894  |
| Db | 258  | TCCGAGCAGGAGATGGTGCAGGTGTTTATTCCTCCGCCAGGCGAGTGGGGCCCATCATCGG   | 317  |
| QY | 895  | CAAGAGGGGCGCAGCATCAAAACAGCTCTCCCGGTTTGCCAGCGCCTCCCATCAAGATTGC   | 954  |
| Db | 318  | CAAGAGGGGCGCAGCATCAAAACAGCTCTCCCGGTTTGCCAGCGCCTCCCATCAAGATTGC   | 377  |
| QY | 955  | ACCACCCGAACACCTGACTCCAAAGTTCGTATGGTTATCATCACTGGACCGCCAGAGGC     | 1014 |
| Db | 378  | ACCACCCGAACACCTGACTCCAAAGTTCGTATGGTTATCATCACTGGACCGCCAGAGGC     | 437  |
| QY | 1015 | CGAATTCAGGCTCAGGGAAGATCTATGCGAACTCAAGGAGGAGNACTCTCTTGGTCC       | 1074 |
| Db | 438  | CGAATTCAGGCTCAGGGAAGATCTATGCGAACTCAAGGAGGAGNACTCTCTTGGTCC       | 497  |
| QY | 1075 | CAAGGAGGAAGTGAAGCTGGAGACCCACATACGTGTGCCAGCATCAGAGCTGGCCGGGT     | 1134 |
| Db | 498  | CAAGGAGGAAGTGAAGCTGGAGACCCACATACGTGTGCCAGCATCAGAGCTGGCCGGGT     | 557  |
| QY | 1135 | CATTGGCAAGGTGGAAAAACGGTGAACGAGTTGACGAATTTGACGGCAGCTCAGGTGGT     | 1194 |
| Db | 558  | CATTGGCAAGGTGGAAAAACGGTGAACGAGTTGACGAATTTGACGGCAGCTCAGGTGGT     | 617  |
| QY | 1195 | AGTACCAGAGAGACAGACCCCTGATGAGAACGACACAGGTTCATCGTGAATAATCATCGGACA | 1254 |
| Db | 618  | AGTACCAGAGAGACAGACCCCTGATGAGAACGACACAGGTTCATCGTGAATAATCATCGGACA | 677  |
| QY | 1255 | TTTTCTATGCCAGTCAGATGGCTCAACGGGAAGATCCGAGACATCTGCGCCAGGTTAAGCA   | 1314 |
| Db | 678  | TTTTCTATGCCAGTCAGATGGCTCAACGGGAAGATCCGAGACATCTGCGCCAGGTTAAGCA   | 737  |
| QY | 1315 | GCAGCATCAGAAAGGACAGAGTAACACGAGGCCACGACGGAGGAAGTGACCGCCCTCC      | 1374 |
| Db | 738  | GCAGCATCAGAAAGGACAGAGTAACACGAGGCCACGACGGAGGAAGTGACCGCCCTCC      | 797  |
| QY | 1375 | CTGTCCCTTTNGAGTCCAGGACAAACAGGGCAGAAATCGAGAGTGTCTCTCCCCGGCAG     | 1434 |
| Db | 798  | CTGTCCCTTCGAGTCCAGGACAAACAGGGCAGAAATCGAGAGTGTCTCTCCCCGGCAG      | 857  |
| QY | 1435 | GCCTGGAATGAGTGGGAATCCGGGACACNTGGCCGGGCTGTAGATCAGGTTTGGCCAC      | 1494 |
| Db | 858  | GCCTGGAATGAGTGGGAATCCGGGACACCTGGCCGGGCTGTAGATCAGGTTTGGCCAC      | 917  |
| QY | 1495 | TTGATTGAGAAAGATGTTCCAGTCAGGAACCTGATCTNTCAGCCCCCAACACCCACCCA     | 1554 |
| Db | 918  | TTGATTGAGAAAGATGTTCCAGTCAGGAACCTGATCTNTCAGCCCCCAACACCCACCCA     | 977  |
| QY | 1555 | ATTGGCCCAACACTGNTNGCCCTCGGGGTGTGAGAAATTNTWAGCGCAAGGCACTTTTAA    | 1614 |
| Db | 978  | ATTGGCCCAACACTGNTNGCCCTCGGGGTGTGAGAAATTNTWAGCGCAAGGCACTTTTAA    | 1037 |
| QY | 1615 | ACGTGGATTGTTTAAAGAGGCTCTCCAGGCCCCACCAAGAGGGTGGATCACACCTCAGTG    | 1674 |
| Db | 1038 | ACGTGGATTGTTTAAAGAGGCTCTCCAGGCCCCACCAAGAGGGTGGATCACACCTCAGTG    | 1697 |
| QY | 1675 | GGAGAGAAAAATAAAATTTCCCTTCAGGTTTTAAAA                            | 1708 |
| Db | 1098 | GGAGAGAAAAATAAAATTTCCCTTCAGGTTTTAAAA                            | 1131 |

RESULT 7  
AAS26566  
ID AAS26566 standard; cDNA; 1129 BP.



Qy 851 GTGAGTGTGTTATCCCGCCAGGAGTGGCGCCATCATCGGAAGAGGGCGAGCAC 910  
Db 241 GTGAGTGTGTTATCCCGCCAGGAGTGGCGCCATCATCGGAAGAGGGCGAGCAC 300  
Qy 911 ATCAAAAGAGTCTCCCGGTTGGCAGCGCTCCATCAAGATTGACCAACCCGAAACACT 970  
Db 301 ATCAAAAGAGTCTCCCGGTTGGCAGCGCTCCATCAAGATTGACCAACCCGAAACACT 360  
Qy 971 GACTCCAAAGTGTGTTATCATCACTGGACCGCCAGAGGCCCAATTCAGGCTCAG 1030  
Db 361 GACTCCAAAGTGTGTTATCATCACTGGACCGCCAGAGGCCCAATTCAGGCTCAG 420  
Qy 1031 GGAAGATCTATGCAAACTCAAGAGGAGAACTTTTGGTCCCAAGAGGAAAGTGAAG 1090  
Db 421 GGAAGATCTATGCAAACTCAAGAGGAGAACTTTTGGTCCCAAGAGGAAAGTGAAG 480  
Qy 1091 CTGAGACCCACATACGTGTGCCAGCATCAGCAGTGGCCGGGTCACTGTCGCAAGGTGGA 1150  
Db 481 CTGAGACCCACATACGTGTGCCAGCATCAGCAGTGGCCGGGTCACTGTCGCAAGGTGGA 540  
Qy 1151 AAAACGGTGAACGAGTTGCAAGATTTTACGCGCAGCTGAGTGGTACCAAGAGACCAG 1210  
Db 541 AAAACGGTGAACGAGTTGCAAGATTTTACGCGCAGCTGAGTGGTACCAAGAGACCAG 600  
Qy 1211 ACCCTCATGAGAACGACCAAGTTCATCGTGAATAATCATCGACATTTCTATGCCAGTCA 1270  
Db 601 ACCCTCATGAGAACGACCAAGTTCATCGTGAATAATCATCGACATTTCTATGCCAGTCA 660  
Qy 1271 ATGCTCAAACGAGATCCGAGACATCTGCGCCAGGTTAAGCAGCAGCATCAGAGGGA 1330  
Db 661 ATGCTCAAACGAGATCCGAGACATCTGCGCCAGGTTAAGCAGCAGCATCAGAGGGA 720  
Qy 1331 CAGAGTAACAGGCGCCAGGACGAGGAAGTGACCGCCCTCCCTGCTCCCTTNGAGTCC 1390  
Db 721 CAGAGTAACAGGCGCCAGGACGAGGAAGTGACCGCCCTCCCTGCTCCCTTNGAGTCC 780  
Qy 1391 AGGACAAACAGGCGAGAAATCGAGATGTGCTCTCCCGGCGAGGCTGAGATGAGTGG 1450  
Db 781 AGGACAAACAGGCGAGAAATCGAGATGTGCTCTCCCGGCGAGGCTGAGATGAGTGG 840  
Qy 1451 GAATCCGGGACACTGCGCGGGCTGTAGATCAGGTTTGGCCACTTCAATGAGAAAGATG 1510  
Db 841 GAATCCGGGACACTGCGCGGGCTGTAGATCAGGTTTGGCCACTTCAATGAGAAAGATG 900  
Qy 1511 TTCCAGTGAGGAACCTGATCTNTCAGCCCAACACCCCAATTTGGCCCAACACTGT 1570  
Db 901 TTCCAGTGAGGAACCTGATCTNTCAGCCCAACACCCCAATTTGGCCCAACACTGT 960  
Qy 1571 NTGCCCTCCGGTGTGAGAAATTTAGCGCAAGCACTTTTAAACGTGATTTTAAA 1630  
Db 961 CTGCCCTCCGGTGTGAGAAATTTAGCGCAAGCACTTTTAAACGTGATTTTAAA 1020  
Qy 1631 GAAGCTCTCCAGGCGCCACCAAGAGGTTGATCACCTCAGTGGGAGGAAATATAAT 1690  
Db 1021 GAAGCTCTCCAGGCGCCACCAAGAGGTTGATCACCTCAGTGGGAGGAAATATAAT 1080  
Qy 1691 TTCCTCAGGTTTAAA 1708  
Db 1081 TTCCTCAGGTTTAAA 1098

RESULT 8  
ABX73907  
ID ABX73907 standard; DNA; 1129 BP.  
XX  
AC ABX73907;  
XX  
DT 18-MAR-2003 (first entry)  
XX  
DE Human novel polynucleotide #735.  
XX  
KW Human; gene; ds; neural disorder; immune system disorder; renal disorder;

KW muscular disorder; respiratory disease; reproductive disorder;  
KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;  
KW hyperproliferative disorder; inflammatory disease; allergic reaction;  
KW blood related disorder; cancer; immunosuppressive; antiinflammatory;  
KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;  
XX haemostatic; antiarteriosclerotic.  
OS Homo sapiens. XX  
XX US2002132753-A1.  
PN 19-SEP-2002. XX  
PD 17-JAN-2001; 2001US-00764864. XX  
PF 31-JAN-2000; 2000US-0179065P. XX  
PR 04-FEB-2000; 2000US-0180628P. XX  
PR 28-JUN-2000; 2000US-0214888P. XX  
PR 07-JUL-2000; 2000US-0216647P. XX  
PR 07-JUL-2000; 2000US-0216880P. XX  
PR 11-JUL-2000; 2000US-0217487P. XX  
PR 11-JUL-2000; 2000US-0217496P. XX  
PR 14-JUL-2000; 2000US-0218290P. XX  
PR 26-JUL-2000; 2000US-0220963P. XX  
PR 26-JUL-2000; 2000US-0220964P. XX  
PR 14-AUG-2000; 2000US-0224518P. XX  
PR 14-AUG-2000; 2000US-0224519P. XX  
PR 14-AUG-2000; 2000US-0225267P. XX  
PR 14-AUG-2000; 2000US-0225268P. XX  
PR 14-AUG-2000; 2000US-0225270P. XX  
PR 14-AUG-2000; 2000US-0225447P. XX  
PR 14-AUG-2000; 2000US-0225757P. XX  
PR 22-AUG-2000; 2000US-0226868P. XX  
PR 30-AUG-2000; 2000US-0228924P. XX  
PR 01-SEP-2000; 2000US-0229287P. XX  
PR 01-SEP-2000; 2000US-0229343P. XX  
PR 01-SEP-2000; 2000US-0229344P. XX  
PR 01-SEP-2000; 2000US-0229345P. XX  
PR 05-SEP-2000; 2000US-0229509P. XX  
PR 05-SEP-2000; 2000US-0229513P. XX  
PR 08-SEP-2000; 2000US-0231413P. XX  
PR 21-SEP-2000; 2000US-0234223P. XX  
PR 21-SEP-2000; 2000US-0234274P. XX  
PR 25-SEP-2000; 2000US-0234997P. XX  
PR 27-SEP-2000; 2000US-0235834P. XX  
PR 29-SEP-2000; 2000US-0236327P. XX  
PR 29-SEP-2000; 2000US-0236367P. XX  
PR 29-SEP-2000; 2000US-0236368P. XX  
PR 29-SEP-2000; 2000US-0236369P. XX  
PR 29-SEP-2000; 2000US-0236370P. XX  
PR 02-OCT-2000; 2000US-0236802P. XX  
PR 02-OCT-2000; 2000US-0237037P. XX  
PR 02-OCT-2000; 2000US-0237038P. XX  
PR 02-OCT-2000; 2000US-0237039P. XX  
PR 02-OCT-2000; 2000US-0237040P. XX  
PR 13-OCT-2000; 2000US-0239935P. XX  
PR 20-OCT-2000; 2000US-0240960P. XX  
PR 20-OCT-2000; 2000US-0241785P. XX  
PR 20-OCT-2000; 2000US-0241809P. XX  
PR 01-NOV-2000; 2000US-0244617P. XX  
PR 17-NOV-2000; 2000US-0249299P. XX  
PR 08-DEC-2000; 2000US-0251856P. XX  
PR 08-DEC-2000; 2000US-0251868P. XX  
PR 08-DEC-2000; 2000US-0251869P. XX  
XX  
XX (ROSE/) ROSEN C A.  
XX (RUBE/) RUBEN S M.  
XX (BARA/) BARASH S C.  
XX  
XX Rosen CA, Ruben SM, Barash SC;  
XX WPI; 2003-147444/14.  
DR

DR P-PSDB; ABUS5647.  
XX  
PT New polypeptides and nucleic acids, useful in gene therapy for treating,  
PT inhibiting or preventing e.g. neural, immune system, muscular,  
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or  
PT renal disorders.  
XX  
PS Claim 1; SEQ ID NO 745; 402pp; English.  
PS  
CC The invention relates to human novel polypeptides and their associated  
CC polynucleotides. The polypeptides and polynucleotides are useful in gene  
CC therapy for treating, inhibiting or preventing neural disorders, immune  
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis  
CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.  
CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,  
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders  
CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left  
CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage  
CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and  
CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and  
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood  
CC related disorders (e.g. thrombosis, atherosclerosis and myocardial  
CC infarction) and cancerous diseases. Sequences ABX73173-ABX74167 represent  
CC human novel polynucleotides of the invention  
XX  
SQ Sequence 1129 BP; 318 A; 296 C; 302 G; 210 T; 0 U; 3 Other;  
Query Match 62.6%; Score 1069.8; DB 7; Length 1129;  
Best Local Similarity 98.7%; Pred. No. 1e-292;  
Matches 1084; Conservative 3; Mismatches 5; Indels 6; Gaps 1;  
QY 617 ACTGTGAAGGGGCGATCGAGAAATGTTGAGGCGGAGGAGGAAATATGAAGAAAGTT 676  
Db 1 ACTGTGAAGGGGCGATCGAGAAATGTTGAGGCGGAGGAGGAAATATGAAGAAAGTT 60  
QY 677 CGGAGGCGCTATGAGATGATGGCTGCCATGAGC-----TCTCACCTGATCCCTGGC 730  
Db 61 CGGAGGCGCTATGAGAAATGATGGCTGCCATGAGCTGAGCTCACCTGATCCCTGGC 120  
QY 731 CTGAACCTGGCTGTAGTGTCTTTCCAGCTTCATCCAGCGAGTCCCGCGCTCC 790  
Db 121 CTGAACCTGGCTGTAGTGTCTTTCCAGCTTCATCCAGCGAGTCCCGCGCTCC 180  
QY 791 AGCAGCGTTACTGGGCTGCTCCCTATAGTCTCTTTATGAGGCTCCGAGGAGATG 850  
Db 181 AGCAGCGTTACTGGGCTGCTCCCTATAGTCTCTTTATGAGGCTCCGAGGAGATG 240  
QY 851 GTCCAGGTGTTTATCCCGCGCCAGGAGTGGGCGCCATCATCGGCAAGAGGGGCGAGC 910  
Db 241 GTCCAGGTGTTTATCCCGCGCCAGGAGTGGGCGCCATCATCGGCAAGAGGGGCGAGC 300  
QY 911 ATCAACAGCTCTCCCGGTTTGCAGGCGCTCCATCAAGATTGCCACCCCGAAACCT 970  
Db 301 ATCAACAGCTCTCCCGGTTTGCAGGCGCTCCATCAAGATTGCCACCCCGAAACCT 360  
QY 971 GACTCCAAAGTTCGTATGTTATCATCATGAGCGCCAGGAGGAGGAGGAGGAGGAGG 1030  
Db 361 GACTCCAAAGTTCGTATGTTATCATCATGAGCGCCAGGAGGAGGAGGAGGAGGAGG 420  
QY 1031 GGAAGATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAGGAGG 1090  
Db 421 GGAAGATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAGGAGG 480  
QY 1091 CTGGAGACCCACATAGTGTGCGAGCATCAGCAGCTGGCGGCTCATTCGCAAGGAGTGA 1150  
Db 481 CTGGAGACCCACATAGTGTGCGAGCATCAGCAGCTGGCGGCTCATTCGCAAGGAGTGA 540  
QY 1151 AAAACCGTGAACGAGTTCAGAAATTTGACGGAGCTGAGGTGGTAGTACCAAGAGACCA 1210  
Db 541 AAAACCGTGAACGAGTTCAGAAATTTGACGGAGCTGAGGTGGTAGTACCAAGAGACCA 600  
QY 1211 ACCCTGTATGAGAACGAGGAGTTCATCGTGAATTCATCGGACATTTCTATCCAGTCA 1270  
Db 1211 ACCCTGTATGAGAACGAGGAGTTCATCGTGAATTCATCGGACATTTCTATCCAGTCA

Db 601 ACCCTGTATGAGAACGAGGAGTTCATCGTGAATTCATCGGACATTTCTATCCAGTCA 660  
QY 1271 ATGGTCTAACGAGGAGTTCATCGGAGATCTCTGCGCCAGGTTAAGCAGCAGCATCAGA 1330  
Db 661 ATGGTCTAACGAGGAGTTCATCGGAGATCTCTGCGCCAGGTTAAGCAGCAGCATCAGA 720  
QY 1331 CAGAGTAAACGAGGAGGAGTTCATCGGAGGAGTTCATCGGAGGAGTTCATCGGAGTTC 1390  
Db 721 CAGAGTAAACGAGGAGGAGTTCATCGGAGGAGTTCATCGGAGGAGTTCATCGGAGTTC 780  
QY 1391 AGGACAAACGAGGAGGAGTTCATCGGAGGAGTTCATCGGAGGAGTTCATCGGAGTTC 1450  
Db 781 AGGACAAACGAGGAGGAGTTCATCGGAGGAGTTCATCGGAGGAGTTCATCGGAGTTC 840  
QY 1451 GAATCCGGGACACNTGGGCGGCTGTAGATCAGGTTTGGCCACTTGTGATTTGAGAAAG 1510  
Db 841 GAATCCGGGACACNTGGGCGGCTGTAGATCAGGTTTGGCCACTTGTGATTTGAGAAAG 900  
QY 1511 TTCCAGTGAAGAACCTTCATCTNTCAGCCCAACACACACCCCAATTTGGCCCAACACT 1570  
Db 901 TTCCAGTGAAGAACCTTCATCTNTCAGCCCAACACACACCCCAATTTGGCCCAACACT 960  
QY 1571 NTGCCCCCTCGGGGTGTCAAGAAATTTAGCGCAAGGAGTTCATTTAAACGTTGATTT 1630  
Db 961 CTGCCCCCTCGGGGTGTCAAGAAATTTAGCGCAAGGAGTTCATTTAAACGTTGATTT 1020  
QY 1631 GAAGCTCTCCAGGCGCCACCAAGAGGTTGATCATCAGTTCAGTTCGGAAGAAATTA 1690  
Db 1021 GAAGCTCTCCAGGCGCCACCAAGAGGTTGATCATCAGTTCAGTTCGGAAGAAATTA 1080  
QY 1691 TTCCCTCAGGTTTAAAA 1708  
Db 1081 TTCCCTCAGGTTTAAAA 1098  
RESULT 9  
AAK91969  
ID AAK91969 standard; cDNA; 833 BP.  
XX  
AC AAK91969;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human cDNA 5'-end sequence, SEQ ID NO: 429.  
XX  
KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.  
XX  
OS Homo sapiens.  
XX  
PN EP1130094-A2.  
XX  
PD 05-SEP-2001.  
XX  
PF 07-JUL-2000; 2000EP-00114089.  
XX  
PR 08-JUL-1999; 98JP-00194486.  
XX  
PR 11-JAN-2000; 2000JP-00118774.  
XX  
PR 02-MAY-2000; 2000JP-00183765.  
XX  
PA (HELI-) HELIX RES INST.  
XX  
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;  
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;  
XX  
XX WPI; 2001-524255/58.  
XX  
PR 930 Primers useful for synthesizing full length cDNA clones and their use  
PT in genetic manipulation.  
XX  
PS Claim 2; SEQ ID NO 429; 1380pp + Sequence Listing; English.  
XX  
CC The invention relates to primers for synthesizing full length cDNA  
CC clones. 930 cDNA molecules encoding a human protein have been isolated





QY 539 GTAGAGCAGATACCGACACAAAATCACCATCTCTCGTTGCAAGACCTTACCTTTAC 598  
XX |||||  
Db 241 GTAGAGCAGATACCGACACAAAATCACCATCTCTCGTTGCAAGACCTTACCTTTAC 300  
QY 599 AACCTTGAGAGACCATCACTGTGTAAGGGGCGCCATCGAGAAATTTGTCAGGGCGGAGCAG 658  
Db 301 AACCTTGAGAGACCATCACTGTGTAAGGGGCGCCATCGAGAAATTTGTCAGGGCGGAGCAG 360  
QY 659 GAAATAATGAGAAGTTGCGAGGCGCTATGAGAAATGATGGTGCATGAGC----- 712  
XX |||||  
Db 361 GAAATAATGAGAAGTTGCGAGGCGCTATGAGAAATGATGGTGCATGAGCCTGAGC 420  
QY 713 TCTCACTGATCTCCCTGGGCTGAACCTGCTGCTGTAGTCTTTTCCAGCTTTCATCCAGC 772  
Db 421 TCTCACTGATCTCCCTGGGCTGAACCTGCTGCTGTAGTCTTTTCCAGCTTTCATCCAGC 480  
QY 773 GCAGTCCCGCGCTCCAGAGGCTTACTGGGCTGTCTCCCTATAGTCTTTTATGAGC 832  
Db 481 GCAGTCCCGCGCTCCAGAGGCTTACTGGGCTGTCTCCCTATAGTCTTTTATGAGC 540  
QY 833 GCTCCGAGCAGGAGATGCTGAGGTGTTATCCCGCCAGGCGAGTGGGCGCATCATC 892  
Db 541 GCTCCGAGCAGGAGATGCTGAGGTGTTATCCCGCCAGGCGAGTGGGCGCATCATC 600  
QY 893 GGCAAGAGGGGCGAGCATCAACAGCTCTCCCGGTTTGCAGCGCTCCATCAAGATT 952  
Db 601 GGCAAGAGGGGCGAGCATCAACAGCTCTCCCGGTTTGCAGCGCTCCATCAAGATT 660  
QY 953 GCACACCCGAAACACCTGACTCCAAAGTTGATGGTTATCATCACTGGA-CCGCCAGA 1011  
Db 661 GCACACCCGAAACACCTGACTCCAAAGTTGATGGTTATCATCACTGGAACCGCCAGA 720  
QY 1012 GCGCCAAFTCAAGGCTCAGGGAAGAATCTATGGCAAACTC-AAGGAGGAGAACTT--CTT 1068  
Db 721 NGCCAAFTCAAGGCTCANGGAAGAATCTATGGCAAACTCAAGGAGGAGAACTTTCTTT 780  
QY 1069 TGGTCCCAAGAGGAGT--GAAGCTGGAGACCCACATAGTGTGCCA 1114  
Db 781 GGTCCCCAAGGAGGAGTGAAGCTTGAGACCCACATACCTGTGCCA 828

RESULT 11  
AAC66035 standard; cDNA; 1740 BP.  
XX AC AAC66035;  
XX 21-FEB-2001 (first entry)  
XX Human lung cancer-associated cDNA antigen L523S.  
XX Lung cancer; therapy; treatment; human; tumor; immunogenic; cytostatic;  
XX vaccine; detection; ss.  
XX Homo sapiens.  
XX WO200061612-A2.  
XX 19-OCT-2000.  
XX 03-APR-2000; 2000WO-US008896.  
XX 02-APR-1999; 99US-00285479.  
XX 17-DEC-1999; 99US-00466396.  
XX 30-DEC-1999; 99US-00476496.  
XX 10-JAN-2000; 2000US-00480884.  
XX 22-FEB-2000; 2000US-00510376.  
XX (CORI-) CORIXA CORP.  
XX Wang T, Fan L;  
XX WPI; 2000-628399/60.  
DR

DR P-PSDB; AAB11365.  
XX Isolated polypeptide comprising an immunogenic portion of a lung tumor  
PT protein is used for detecting and monitoring progression of lung cancer  
PT in a patient.  
XX Claim 1a; Page 258-259; 261pp; English.  
XX This invention describes a novel isolated polypeptide (I) which  
CC comprising an immunogenic portion of a lung tumor protein or variant (P2)  
CC which have cytostatic activity. The polypeptides and polynucleotides are  
CC used in compositions and vaccines to inhibit the development of cancer,  
CC especially lung cancer, in a patient. Methods described in the invention  
CC can be used to monitor the progression of a cancer by carrying out the  
CC detection at subsequent time points and comparing the results from the  
CC different time points. CD4+ and/or CD8+ T-cells isolated from a patient  
CC are treated with P2, polynucleotides encoding P2 or antigen presenting  
CC cells expressing P2 and then administered to the patient to inhibit  
CC development of cancer  
XX  
SQ Sequence 1740 BP; 526 A; 406 C; 417 G; 391 T; 0 U; 0 Other;  
Query Match 34.1%; Score 582.2; DB 3; Length 1740;  
Best Local Similarity 66.2%; Pred. No. 2.8e-154;  
Matches 858; Conservative 0; Mismatches 433; Indels 6; Gaps 1;  
QY 37 AGCCATCATGAGCTGAATGGCCACCACTGTGGAGAACCATGCCCTGAGAGGTCCTCTACAT 96  
Db 402 AGCACTAGACAACTGAATGGATTTCACTTGAAGATTTTCACTTGAAGATGAGCTATAT 461  
QY 97 CCCGATGAGCAGATGAGCAGGAGCCTGAGAAATGGCGCGGAGGGGCTTTGGCTCTCG 156  
Db 462 CCCTGATGAACGGCGCGCCAGCAAAACCCCTTGAGCAGAGCCCGAGGTCCCGGGGCT 521  
QY 157 GGGTCAGCCCCCAGGCGCTCACCTGTGGCAGCGGGGCCCCAGCAGAGCAGCAAGT 216  
Db 522 TGGGAGAGAGGGGCTCCTCAAGCAGGGGTCTCCAGGATCCGTATCCAGCAGAAACCATG 581  
QY 217 GGACATCCCGCTTCGGCTCCTGGTCCCAACCACTGATGGGTGCTATTTTGGCAAGGA 276  
Db 582 TGAATTTGCTCTGCGCTGCTGGTTCACCACTTGTGGAGCCATCATAGAAAGA 641  
QY 277 GGGGGCCACCATCCGCAACATCAAAACAGACCCAGTCCCAAGATGAGCTGATAGAA 336  
Db 642 AGGTGCCACCATTCGGAAACATCACCAGACAGACCCAGTCTAAAATCGATGTGCCACG 701  
QY 337 GGAGAACGAGGTGCAGCTGAAAGAGCCATCAGTGTGCACCTCCACCCCTGAGGGCTGCTC 396  
Db 702 AGAAATGGGGGCTGCTGAGAGTCAATTAATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 761  
QY 397 CTCGCTTGTAGATGATCTTTGGAGATPATGCAATAAGAGGCTAAGGACACCAAAACGGC 456  
Db 762 TCGGCTTGTAAATCTATCTGGAGATTTATGAGAGGAGCTCAAGATATAAAATTCAC 821  
QY 457 TGACGAGGTCCCTGAGATCTCGCCCAATAAATTTGTAGGGGCTCTCATTGGCAA 516  
Db 822 AGAAGAGATCCCTTGAAGATTTAGTCTATAAATCTTTGTCAGCTCTTATTTGGTAA 881  
QY 517 GGAAGGACGAACTGAGAGAGGTAGAGCAAGATACCGAGACAAATAATCACCATCTCTCTC 576  
Db 882 AGAAGGAGAAATCTTAAAAAATTTGAGCAAGACACAGACACTAAATCAGATATCTCC 941  
QY 577 GTTGCAGACCTTACCTTTTACACCTTGAGAGAGCCTCATCTGTGAGGGGCGCATCGA 636  
Db 942 ATTGCAGAAATGACGCTGTATATCCAGAACCGCACTATTACAGTTAAAGGCAATGTGA 1001  
QY 637 GAAATGTTGCGGGCGGAGCAGGAGAAATTAATGAGAAAGTTTCGGGAGGCTTATGAGAA 696  
Db 1002 GACATGTGCCAAAGCTGAGGAGGAGATCATGAGAAATCAGGAGTCTTATGAAATGA 1061  
QY 697 TGTGGCTCCCATGA-----GCTCTACCTGATCCCTGGCTGACCTGGCTGCTTAGG 750  
Db 1062 TATTGCTTCTATGAATCTCAAGCACATTTAAATTCCTGGATTAATCTGAAACGCGCTTGGG 1121



Qy 751 TCITTTCCAGCTTCATCCAGCGAGTCCCGCCCTCCAGCAGCGTTACTGGGGTGC 810  
Db |||||  
Qy 1122 TCTGTTCACCCACTTCAGGATGCCACCTCCACCTCAGGGGCCCTTCAGCCATGAC 1181  
Db |||||  
Qy 811 TCCCTATAGCTCCTTTATGCAAGCTCCGAGCAGGAGTGGTCAGGTGTTTATCCCGC 870  
Db |||||  
Qy 1182 TCCCTCCATCCGAGTTGAGCAATCAGAAACGAGACTGTTCTATCTGTTATCCGAGC 1241  
Db |||||  
Qy 871 CCAGGAGTGGCGCCCATCATCGGCAAGAGGGGCGAGCAGATCAACAGAGCTCTCCCGGTT 930  
Db |||||  
Qy 1242 TCTATCAGTGGTCCCATCATCGGCAAGCAGGCGCAGCAGATCAAGCAGCTTCTCGCTT 1301  
Db |||||  
Qy 931 TCCAGCGCTCCATCAAGATTGACACCCGCAACCTGACCTCCAAAGTTCGTATGCT 990  
Db |||||  
Qy 1302 TCGTGGAGCTTCAATTAGATTGCTCCAGCGGAACCAAGATCTAAAGTGAGGATGGT 1361  
Db |||||  
Qy 991 TATCATCACTGAGCGCGCAGAGGCCCAATTCAAGGCTCAGGGAAGAATCTATGGCAAACT 1050  
Db |||||  
Qy 1362 GATTATCACTGAGCAGCAGAGGCTCAGTTCAGGCTCAGGGAAGAATTTATGGAAAAAT 1421  
Db |||||  
Qy 1051 CAAGGAGGAGAACTCTTTGTTGCCAAGAGGAGGAAGTGAAGTGGAGACCCACATACGTGT 1110  
Db |||||  
Qy 1422 TAAAGAGAAAACTTTGTTAGTCTTAAAGAGAGGTGAACTTGAAGCTCATATCAGAGT 1481  
Db |||||  
Qy 1111 GCCAGCATCAGAGCTGCGCGGTCAATTGGCAAGGTGGAAGGCTGAACGAGTTGCA 1170  
Db |||||  
Qy 1482 GCCATCTCTTGTCTGCGCAGAGTTATTGGAAAGGAGGCAAAACGGTGAATGAATCA 1541  
Db |||||  
Qy 1171 GAAATTGACGGCAGCTGAGGTGGTAGTACCAGAGACAGACCCCTGATGAGAACGACCA 1230  
Db |||||  
Qy 1542 GAAATTGTCAGTGACAGAGTTGTTGCTCCGTCGACAGACACTGATGAGATGACCA 1601  
Db |||||  
Qy 1231 GGTCTCTGTAATATCGACACATTTCTATGCCAGTCAGATGCTCAGCGAGATCCG 1290  
Db |||||  
Qy 1602 AGTGGTTCTCAAAATAACTGCTCACTTCTATGCTTGCAGGTTGCCAGAGAAAAATCA 1661  
Db |||||  
Qy 1291 AGACATCTGCGCCAGGTTAAGCAGCAGCATCAGAG 1327  
Db |||||  
Qy 1662 GGAAATTCGACTAGGTAAAGCAGCACCACACAG 1698  
Db |||||

RESULT 12  
ABL49254  
ID ABL49254 standard; cDNA; 1740 BP.  
XX ABL49254;  
AC ABL49254;  
XX  
XX  
XX 01-MAY-2002 (first entry)  
XX Human lung tumour L523S cDNA sequence SEQ ID NO:347.  
XX Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;  
XX immune response; ss.  
XX Homo sapiens.  
XX OS  
XX WO200200174-A2.  
XX PN  
XX  
XX PD 03-JAN-2002.  
XX PF  
XX PF 28-JUN-2001; 2001WO-US021065.  
XX PR  
XX PR 28-JUN-2000; 2000US-00606421.  
XX PR 02-AUG-2000; 2000US-00630940.  
XX PR 21-AUG-2000; 2000US-00643597.  
XX PR 15-SEP-2000; 2000US-00662786.  
XX PR 09-OCT-2000; 2000US-00685696.  
XX PR 12-DEC-2000; 2000US-00735705.  
XX PR 07-MAY-2001; 2001US-00850716.  
XX  
XX (CORI-) CORIXA CORP.

PI Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;  
PI McNeill PD, Fanger N, Retter MW, Marnerakis M, Fanger GR;  
XX Vedvick TS, Carter D, Watanabe Y, Peckham DW;  
DR WPI; 2002-090513/12.  
DR P-PSDB; ABB74957.  
XX Polynucleotides encoding lung tumor polypeptides, useful for treating  
PT lung cancer or stimulating an immune response.  
XX Example 2; Page 330; 374pp; English.  
XX The present invention describes human lung tumour proteins. Human lung  
CC tumour proteins and polynucleotides have cytostatic and immunostimulant  
CC activities, and can be used in vaccine production. Compositions  
CC comprising the lung tumour proteins, polynucleotides, antibodies, fusion  
CC proteins, T cell populations, or antigen presenting cells that express  
CC the lung tumour proteins are useful for treating lung cancer or  
CC stimulating an immune response. ABL48959 to ABL49300 and ABB74946 to  
CC ABB75070 represent sequences used in the exemplification of the present  
XX invention.  
SQ Sequence 1740 BP; 526 A; 406 C; 417 G; 391 T; 0 U; 0 Other;  
Query Match 34.1%; Score 582.2; DB 6; Length 1740;  
Best Local Similarity 66.2%; Pred. No. 2.8e-154;  
Matches 858; Conservative 0; Mismatches 433; Indels 6; Gaps 1;  
Qy 37 AGCCATCATGAGCTGAATGGCCACCAGTTGGAGAACCATGCTTCAAGGTCTCTACAT 96  
Db 402 AGCCTAGACAACTGAATGGATTCAGTTAGAGATTTTCACTTGAAGTAGCTTATAT 461  
Qy 97 CCCCGATGAGCAGATGACAGGGACCTGAGAATGGCGCCGAGGGGGCTTTGGCTCTCG 156  
Db 462 CCTGATGAAACGGCGCCCGCCAGCAAAACCTCTGACAGAGCCCGCGGGGCT 521  
Qy 157 GGTCTAGCCCGCCAGGGCTCACTGTGGCAGGGGGGCCAGCCAGCAGCAGCAAGT 216  
Db 522 TGGCGCAGAGGGGCTCCTCAAGGCGAGGGGTCTCAGGATCCGTATCCAGCAGAAACATG 581  
Qy 217 GGACATCCCTCTCGGCTCTGCTGTCGCCACCATGATGTGGTGCCATTATTGGCAAGGA 276  
Db 582 TGAATTTGCTCTGCGCTCTGCTGTTCCACCCCAATTTGTGGAGCCATCATAGAAAAAGA 641  
Qy 277 GGGGGCCACCATCCGCAACATCAAAAACAGACCCAGTCCAAGATAGAGTGCATAGGAA 336  
Db 642 AGGTGCCACCATTCGGAACATCACCAACAGACCCAGTCTAAATCGATGCCACCGTAA 701  
Qy 337 GGAGACGAGGTGCAGCTGAAAAGCCATCAGTGTGCATCCACCCCTGAGGGCTGCTC 396  
Db 702 AGAAATGGCGGGGCTGCTGAGAAAGTCGATTACTCTCTCTACTCTGAGGACCTC 761  
Qy 397 CTCGGCTTGAAGATCATCTTGAGATTATGCAATAAGAGGCTAAGGACACCAAAACGGC 456  
Db 762 TGGCGCTTGAAGTCTATCTTGAGATTATGCAATAGGAGCTCAAGATATAAATTCAC 821  
Qy 457 TGACGAGGTTCCCTGAAAGATCCTGGCCCAATAATACTTTGTAGGGGCTCTCATGGCAA 516  
Db 822 AGAAGAGATCCCTTGAAGATTTTAGCTATAATAAATTTGTGGAGCTTTATTGGTAA 881  
Qy 517 GGAGGACGGAACCTGGAAGGTGAGGCAAGATACCGAGACAAAATCACCATCTCCTC 576  
Db 882 AGAAGGAGAAATCTTAAAAAATTTGAGCAAGACAGACACTTAATCAAGATATCTCC 941  
Qy 577 GTTGCAGAGCTTTACCTTTTACAACTTGAGAGGACCATCATCTGTGAAGGGGCCATCGA 636  
Db 942 ATTGCAGGAATTTGACGCTGTATAATCCAGAACGCACTATTACAGTTTAAAGGCAATGTGA 1001  
Qy 637 GAATTGTTGAGGGCCGAGCAGGAATAATGAGAAAGTTCGGGAGGCTTATCAGATGA 696  
Db 1002 GACATGTGCCAAAGCTGAGGAGGAGATCATGAAGAAAATCAGGAGGTCTTATGAATAATGA 1061  
Qy 697 TGTGGCTGCCATGA-----GCTCTCACCTGATCCCTGGCCTGAACCTGCTGTGTGAG 750

Db 1062 TATTGCTTCTATGAATCTTCAAGCAATTAATTCCTGGATTAATCTGAACGCTGGG 1121  
QY 751 TCTTTTCCAGCTTTCATCCAGCGAGTCCCGCGCTCCAGCGAGCGTACTAGGGCTGC 810  
Db 1122 TCTGTTCCACCACTTTCAGGATCCACCTCCACCTCAGGCGCCCTTCAGGCATGAC 1181  
QY 811 TCCCTATAGCTCTTTATGAGCTCCCGAGCAGGAGTGTGTCAGTGTGTTATCCCGC 870  
Db 1182 TCTCCTTACCGCAGTTTGAAGCAATCAGAAACGAGACTGTTATCTGTTATCCCGC 1241  
QY 871 CCAGCAGTGGCGCATCATCCGCAAGAGGGGAGCAGACATCAACAGCTCTCCCGTT 930  
Db 1242 TCTATCAGTCGTGCTCATCTCGCAAGAGGGGAGCAGACATCAAGCAGCTTCTCGCT 1301  
QY 931 TCCAGCGCTTCCATCAAGATTGCAACCCCGAAGCACTGACTCCAAAGTTCGTATGTT 990  
Db 1302 TCTCGAGCTTCAATTAAGATTGCTCCAGCGAAGCAGATGCTAAAGTGAGATGTT 1361  
QY 991 TATCATCTGACCGCCAGAGGCCCAATTCAGGCTCAGGGAAGATCTATGCGAACT 1050  
Db 1362 GATTATCACTGACCAACAGAGGCTCAGTTCAAGGCTCAGGGAAGATTTATGAAAAAT 1421  
QY 1051 CAAGAGAGAGAACTTCTTTGTTCCAGAGGAGGAGTGAAGTGGAGACCCACATACGTT 1110  
Db 1422 TAAAGAGAAACTTTGTTAGTCTTAAAGAGAGGTGAACTTGAAGTCAATACAGT 1481  
QY 1111 GCCAGATCAGAGTGGCCCGGCTATTGGCAAGGTTGGAAAAACGGTGAACGATGCA 1170  
Db 1482 GCCATCTTGTCTGTCAGAGTTATTGGAAAAAGGAGGCAAAACGGTGAATGAATCA 1541  
QY 1171 GAAATTGACGAGCTGAGTGTAGTACCAAGAGACAGACCCCTGATGAGACGACCA 1230  
Db 1542 GAAATTGCAAGTGCAGAAAGTTGTGCTCTGTCGACAGACCTGATGAGATGACCA 1601  
QY 1231 GGTATCTGAAATCATCGGACATTTCTATGCGAGTCAGATGGCTCAAGGGAAGATCCG 1290  
Db 1602 AGTGGTTGCAAAATCACTGCTACTTCTATGCTGCGAGTTGCCAGTTGCCAGAAAAATCA 1661  
QY 1291 AGACATCTGCGCCAGGTGAGCAGCAGCATCAGAA 1327  
Db 1662 GGAATTTCTGACTCAGGTAAAGCAGCACCACAAACAG 1698

RESULT 13  
ABQ92440

ID ABQ92440 standard; cDNA; 1740 BP.

XX AC ABQ92440;

XX DT 07-OCT-2002 (first entry)

XX DE Human lung cancer associated cDNA sequence SEQ ID NO:347.

XX KW Human; lung cancer; lung tumour; cytostatic; gene therapy; vaccine; gene;

XX KW ss.

XX OS Homo sapiens.

XX PN WO200247534-A2.

XX PD 20-JUN-2002.

XX PF 30-NOV-2001; 2001WO-0547576.

XX PR 12-DEC-2000; 2000US-00735705.

XX PR 07-MAY-2001; 2001US-00850716.

XX PR 28-JUN-2001; 2001US-00897778.

XX PR (CORI-) CORIXA CORP.

XX PI Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;

PI McNeill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedvick TS;

PI Carter D, Watanabe Y, Peckham DW, Cai F, Foy TW;

XX WPI: 2002-583465/62.

XX P-PSDB; ABP61917.

XX Novel lung carcinoma polynucleotide sequences and polypeptides encoded by the polynucleotides, useful in pharmaceutical compositions such as vaccines and as markers to indicate the presence of lung cancer.

XX Claim 1; Page 337; 381pp; English.

XX The present invention describes isolated human lung carcinoma polynucleotides (I) and polypeptides (II). (I) and (II) have cytostatic activity, and can be used in gene therapy and in vaccines. Compositions comprising (I) or (II) can be used for stimulating an immune response in a patient and for treating lung cancer in a patient. Oligonucleotides of (I) can be used for detecting the presence of a cancer in a patient, by obtaining a biological sample from the patient, contacting the biological sample with the oligonucleotide, detecting in the sample, an amount of polynucleotide that hybridises to the oligonucleotide and comparing the amount of polynucleotide that hybridises to the oligonucleotide to a predetermined cut-off value, and determining the presence of a cancer in the patient. (I) and (II) are useful in pharmaceutical compositions, e.g. vaccines. (I) is useful as a marker to indicate the presence or absence of a cancer such as lung cancer. ABQ92145 to ABQ92486 and ABP61866 to ABP61992 represent sequences used in the exemplification of the present invention

XX SQ Sequence 1740 BP; 526 A; 406 C; 417 G; 391 T; 0 U; 0 Other;

Query Match 34.1%; Score 582.2; DB 6; Length 1740;

Best local Similarity 66.2%; Pred. No. 2.8e-154;

Matches 858; Conservative 0; Mismatches 433; Indels 6; Gaps 1;

QY 37 AGCCATCATGAGCTGAATGGCCAGCTTGGAGAACCATGCCCTGAAGGTCCTCATAC 96

Db 402 ASCACTAGACAACTGAATGGATTTCACTTAGAGAAATTCACCTTGAAGTAGCCTATAT 461

QY 97 CCCGATGAGCAGATAGCACAGGACCTGAGAAATGGCGCGCAGGGGCTTTGGCTCTCG 156

Db 462 CCTGATGAAACGGCGCGCCAGCAAAACCCCTTGCAGCAGCCCGAGGTCGCCGGGCT 521

QY 157 GGGTCAGCCCGCCAGGGCTCACCTGTGGCAGCGGGGCCCCAGCAGCAGCAGCAGT 216

Db 522 TGGCAGAGGGGCTCTCAAGCAGGGGTCTCCAGGATCCGATATCCAGCAGAAACCATG 581

QY 217 GGACATCCCGCTTGGCTCTCGTCCGCCACCCAGATATGTTGGTGCCATATTGGCAAGGA 276

Db 582 TGAATTTGGCTCTGCGCTGCTGGTTCCGCCCAATTTTGGAGCCATCATAGGAAAGA 641

QY 277 GGGGCGCCACCATCGCAACATCACAACACAGACCCAGTCCAGATAGACGTGCATAGAA 336

Db 642 AGGTGCCACCATTCGGAACATCACAACACAGACCCAGTCTAAATCGATGTCCACCGTAA 701

QY 337 GGAGAACCCAGTGCAGCTGAAAGGCCATCAGTGTGCATCTCCACCCCTGAGGGCTGCTC 396

Db 702 AGAAATCGGGGGCTGCTGAGAAAGTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 761

QY 397 CTCGCTTTGTAAGATGATCTTGGAGATATGATATAGAGGCTAAGACACCAACAGGC 456

Db 762 TCGGCTTTGTAAGTCTATTCTGGAGATTTATGATATAGGAGCTCAAGATATAAAATTCAC 821

QY 457 TGACGAGTTCCCTGGAAGATCTTGGGCCCAATAAATTTGTAGGGGCTCTCATTTGCA 516

Db 822 AGAGAGATCCCTTGAAGATTTTAGCTCAATAAATTTTGTGGAGCTCTTATTTGGTAA 881

QY 517 GGAGGACGGAACCTGGAAGAGGTAGAGCAGATACCGAGACAAAATACCATCTCTCTC 576

Db 882 AGAAGGAGAAATCTTAAAAAATTTGAGCAAGACACAGACACTAAAAATCATCTATCTCC 941

QY 577 GTTGAAGACCTTTACCTTTTCAACCCCTGAGAGGACCATCATCTGTAAGGGGGCCATCGA 636

Db 942 ATTGACGAAATTTGACGCTGTATATATCCAGAACGCACTATTACAGTTAAAGCAATGTGA 1001

|    |      |  |      |
|----|------|--|------|
| Qy | 637  | GAATTTGTCAGGGCCGAGCAGGAAATAATGAAGAAAGTTTCGGAGGCCCTATCAGAAATGA    | 696  |
| Db | 1002 | GACATGTGCCAAAGCTGAGGAGGAGATCATGAAGAAATCAGGAGTCTTATGAAATGA        | 1061 |
| Qy | 697  | TGTGGTCCCATGA-----GCTCTCACCTGATCCCTGGCCTGAACCTGCTCCTTAGG         | 750  |
| Db | 1062 | TATTTGCTTCATGAATCTTCAAGACACATTTAAATTCCTGGAAATAATCTGAACCCCTTGGG   | 1121 |
| Qy | 751  | TCTTTTCCCAAGCTTCATCCAGCGCAGTCCCGCCGCCCTCCACAGCAGCGTTACTTGGGGCTGC | 810  |
| Db | 1122 | TCTGTTCCCAACCCATTCAGGATGCCACCTCCCACTCAGGGCCOCITTCAGCCATGAC       | 1181 |
| Qy | 811  | TCCCTATAGTCTCTTTATGCAAGGCTCCGAGCAGGAGATGGTCAGAGTGTTTATCCCGC      | 870  |
| Db | 1182 | TCTTCCCTACCCGAGTTTGAGCAATCAGAAACGGAGACTGTTTCATCTGTTTATCCGACG     | 1241 |
| Qy | 871  | CCAGGCAGTGGCGCCCATCTCGCAAGAGGGGCGACACATCAACACAGCTCTCCCGGTT       | 930  |
| Db | 1242 | TCTATCAGTCGTGCCATCATCGCAAGCAGGGCCAGCACATCAGCAGCTTCTCGCTT         | 1301 |
| Qy | 931  | TGCCAGGGCCCTCCATCAAGATTGCACCAACCGCAACCTCACTTCCAAAGTTCGTATGGT     | 990  |
| Db | 1302 | TGCTGGAGCTTCAATTAAAGATTGCTCAGCGGAAGCACCAGATGCTAAAGTGAGGATGGT     | 1361 |
| Qy | 991  | TATCATCATGTGACCGCCAGAGGCCCAATTCAGGCTCAGGGAAGAACTCTATGGCAACT      | 1050 |
| Db | 1362 | GAATTATCACTGGACCACGAGGGCTCAGTTCAAGGCTCAGGGAAGAAATTTATGGAAAAAT    | 1421 |
| Qy | 1051 | CAGGAGGAGAACTTTTGTGTCACAGGAGGAAGTGAAGCTGGAGACCCACACATACGTGT      | 1110 |
| Db | 1422 | TAAAGAAGAAACHTTGTAGTCTCTAAGAAGAGTGAACTTGAAGCTCATATCAGAGT         | 1481 |
| Qy | 1111 | GCCAGCATCAGCAGCTGCCCGGGTCATTGGCAAAAGGTGGAAAAACGGTGAACCGATTGCA    | 1170 |
| Db | 1482 | GCCATCCTTTGCTGCTGCGCAGTTATTGGAAAAAGGAGGCAAAACGGTGAATGAACTCATCA   | 1541 |
| Qy | 1171 | GAATTTGACGGCAGCTGAGTGGTAGTACCAAGAGACGAGACCCCTCATGAGACACCA        | 1230 |
| Db | 1542 | GAATTTGTCAAGTGCAGAAGTTGTTGTCCCTCGTGAACGAGACCTCATGAGAAATGACCA     | 1601 |
| Qy | 1231 | GGTCATCGTGAATAATCATCGACACATTTCTATGCCAGTCAGATGCTCAACGGAAAGATCCG   | 1290 |
| Db | 1602 | AGTGGTTGTCAAATAACTGTGCTACTTCTATGCTTGCAGGTTGCCACGAGAAAAATCA       | 1661 |
| Qy | 1291 | AGACATCTTGCCCGAGGTTAAGACGACGATCAGAAG                             | 1327 |
| Db | 1662 | GGAAATTCCTGACTCAGGTAAGCAGCAGCAACCAACAG                           | 1698 |

## RESULT 14

ABL49299  
ID ABL49299 standard: cDNA: 1743 BP.

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

AC ABL49299;

































DT 01-MAY-2002 (first entry)

XX DE Human Jung Filmour 15

XX  
DE  
HUMAN TUNG LU

Human: lung tumor

KW immune response

XX

OS Homo sapiens.

[illegible]



QY 577 GTTGCAAGACCTTACCCCTTTACACCTTGAGAGGACCATCACTGTGAAGGGGCCATCGA 636  
Db |||||  
942 ATTGACGAATTGACGCTGTATAATCCAGAACGCACTATTACAGTTAAAGGCAATGTGA 1001  
QY 637 GAATTGTTGACGGCCGAGCAGGAAATAATGAAGAAAGTTGCGGAGSCCTATGAGATGA 696  
Db |||||  
1002 GACATGTGCCAAAGCTGAGGAGGAGATCATGAAGAAATCAGGGAGTCTTTATGAAATGA 1061  
QY 597 TGTGGCTGCCATGA-----GCTCTCACTGATCCCTGGCTGAACCTGGCTGCTGTAGG 750  
Db |||||  
1062 TATTGCTTCTATGAATCTTCAGACACATTTAATTCCTGGATTAAATCTGAAACGCTTGGG 1121  
QY 751 TCTTTTCCAGCTTCATCCAGCGCAGTCCCGCCGCCCTCCAGCAGCGTTACTGGGGCTGC 810  
Db |||||  
1122 TCTGTTCACCCCACTTCAGGGATGCCACCTCCACCTCAGGGCCCCCTTCAGCCATGAC 1181  
QY 811 TCCCTATAGTCTCTTTATGACAGGCTCCGAGCAGGAGATGGTGAGGTGTTTATCCCGC 870  
Db |||||  
1182 TCCTCCCTACCCGAGTTTGAGCAATCAGAAACGAGACTGTTCTCTGTTTATCCGAGC 1241  
QY 871 CCAGGCAGTGGCGCCCATCATCGCAAGAGGGCCAGCACATCAAAACAGCTCTCCCGGTT 930  
Db |||||  
1242 TCTATCAGTGGTGCCCATCATCGCAAGAGGGCCAGCACATCAAGCAGCTTCTCGCTT 1301  
QY 931 TGCCAGCGCTCCATCAAGATTGCACACCGAAACACCTGACTCCAAAGTTCGTATGCT 990  
Db |||||  
1302 TGCTGGAGCTTCAATTAAGATTGCTCCAGCGGAAGCACCAGATGCTAAAGTGAGGATGCT 1361  
QY 991 TATCATCACTGGACCGCCAGAGGCCCAATTCAAGGCTCAGGGAAGAATCTATGGCAAACT 1050  
Db |||||  
1362 GATTATCACTGGACCAACAGAGGCTCAGTTCAAGGCTCAGGGAAGAATTTATGGAAAAAT 1421  
QY 1051 CAAGGAGAGAACTTCTTTGTTCCAGAGGGAAGTGAAGCTGGAGACCCACATACGTTG 1110  
Db |||||  
1422 TAAAGAGAAAACTTTGTTAGTCTTAAAGAGAGGTGAACTTTGAAGCTCATATCAGAGT 1481  
QY 1111 GCCAGCATCAGCAGCTGCGGGTCAATTGGCAAGGTGGAAGAAACGGTGAACGAGTTGCA 1170  
Db |||||  
1482 GCCATCCTTTGCTGTCAGAGTTATTGGAAAGGAGGCAAAACGGTGAATGAATCA 1541  
QY 1171 GAAATTGACGGCAGCTGAGGTGGTAGTACCAGAGACAGACCCCTGATGAGAACGACCA 1230  
Db |||||  
1542 GAAATTGTCAAGTGCAGAAAGTTGTGTCCTCGTGACCCAGACACCTGATGAGAAATGACCA 1601  
QY 1231 GGTCAATCGTGAATATCGACATTTCTATGCCAGTCAAGTGGCTCAACGGAGATCCG 1290  
Db |||||  
1602 AGTGGTTGTCAAAATAAATGCTGCTCTATGCTTGCAGGTGCCCCAGAGAAAAATTC 1661  
QY 1291 AGACATCTCTGGCCAGGTTAAGCAGCAGCATCAGAAG 1327  
Db |||||  
1662 GGAATTTCTGACTCAGGTAAAGCAGCACCACACACAG 1698

Search completed: July 13, 2004, 18:31:39  
Job time : 707 secs



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: July 13, 2004, 12:03:32 ; Search time 30 Seconds  
(without alignments)

10953.008 Million cell updates/sec

Title: US-09-270-437D-5

Perfect score: 3110

Sequence: 1 agggagcgtgcgcacccgc.....atttccttcagggttttaaaa 1708

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 566732

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p,model DEV=xlip  
-Q=/cgn21/USPTO\_spool\_P/US09270437/runat\_13072004\_121922\_9618/app\_query.fasta\_1.1863  
-DB=PIR\_78 -QMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09270437 @CNG 1.1 44 @runat\_13072004\_121922\_9618 -NCPU=6 -ICPU=3  
-NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSFBLOCKS=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PIR\_78:\*  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 451.5 | 14.5        | 854    | 2     | T23837      |
| 2          | 265   | 8.5         | 621    | 2     | D96554      |
| 3          | 256.5 | 8.2         | 398    | 2     | T41600      |
| 4          | 251   | 8.1         | 479    | 2     | C86275      |
| 5          | 248   | 8.0         | 768    | 2     | T27855      |
| 6          | 244   | 7.8         | 644    | 2     | A53184      |
| 7          | 240.5 | 7.7         | 680    | 2     | T25832      |
| 8          | 235.5 | 7.6         | 632    | 2     | T02627      |
| 9          | 234.5 | 7.5         | 588    | 2     | T49962      |
| 10         | 232   | 7.5         | 569    | 2     | T19216      |
| 11         | 232   | 7.5         | 611    | 2     | T19217      |
| 12         | 221.5 | 7.1         | 510    | 2     | T38489      |
| 13         | 218.5 | 7.0         | 356    | 2     | S58529      |
| 14         | 217.5 | 7.0         | 396    | 2     | S41224      |

|    |       |     |      |   |        |                     |
|----|-------|-----|------|---|--------|---------------------|
| 15 | 214   | 6.9 | 313  | 2 | T48439 | probable RNA-bind   |
| 16 | 209   | 6.7 | 397  | 2 | T30168 | hypothetical prote  |
| 17 | 209   | 6.7 | 1268 | 2 | A44125 | high density lipop  |
| 18 | 203.5 | 6.5 | 1270 | 2 | S23464 | vigilin - chicken   |
| 19 | 202.5 | 6.5 | 846  | 2 | T04533 | hypothetical prote  |
| 20 | 199   | 6.4 | 365  | 2 | S42471 | hnRNP protein E2 -  |
| 21 | 197.5 | 6.4 | 463  | 2 | S41495 | dc stretch-binding  |
| 22 | 197   | 6.3 | 464  | 2 | S43363 | transformation upr  |
| 23 | 194   | 6.2 | 362  | 2 | S78515 | single-stranded nu  |
| 24 | 194   | 6.2 | 1279 | 2 | T41389 | rna binding protei  |
| 25 | 193   | 6.2 | 464  | 2 | A54143 | kappa-B motif-bind  |
| 26 | 191   | 6.1 | 413  | 2 | S46109 | hnRNP complex prot  |
| 27 | 190   | 6.1 | 1198 | 2 | B88279 | protein C08H9, 2 li |
| 28 | 190   | 6.1 | 1220 | 2 | T19117 | hypothetical prote  |
| 29 | 182.5 | 5.9 | 1222 | 2 | S56030 | SCF160 protein - y  |
| 30 | 178   | 5.7 | 381  | 2 | S45766 | hypothetical prote  |
| 31 | 164   | 5.2 | 784  | 2 | JQ0317 | hypothetical 82K p  |
| 32 | 163   | 5.2 | 806  | 2 | T13690 | hypothetical prote  |
| 33 | 159   | 5.1 | 389  | 2 | C86460 | FlA2.18 protein -   |
| 34 | 156.5 | 5.0 | 891  | 2 | G84693 | probable proline-r  |
| 35 | 155.5 | 5.0 | 228  | 2 | S42933 | hypothetical prote  |
| 36 | 155.5 | 5.0 | 255  | 2 | JQ0320 | hypothetical 24.7K  |
| 37 | 152   | 4.9 | 1283 | 2 | T49692 | related to SCF160   |
| 38 | 151.5 | 4.9 | 495  | 2 | T04255 | hypothetical prote  |
| 39 | 148   | 4.8 | 1872 | 2 | S36152 | MHC class III hist  |
| 40 | 146   | 4.7 | 1106 | 2 | JQ0405 | hypothetical 119.5  |
| 41 | 144   | 4.6 | 748  | 2 | T84011 | hypothetical prote  |
| 42 | 144   | 4.6 | 1870 | 2 | S37671 | MHC class III hist  |
| 43 | 143.5 | 4.6 | 2187 | 2 | T30826 | nascent polypeptid  |
| 44 | 143   | 4.6 | 648  | 2 | S04832 | glutenin high mole  |
| 45 | 142   | 4.6 | 2142 | 2 | B35098 | MHC class III hist  |

#### ALIGNMENTS

##### RESULT 1

T23837

hypothetical protein M88.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Data: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: T23837

R:Submitted to the EMBL Data Library, June 1994

A:Reference number: Z19806

A:Accession: T23837

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-854 <WIL>

A:Cross-references: EMBL:T34802; PIDN:CAA84338.1; GSPDB:GN000021; CESP:M88.5

A:Experimental source: clone M88

C:Genetics:

A:Gene: CESP:M88.5

A:Map position: 3

A:Introns: 39/3; 96/1; 133/3; 255/3; 307/3; 395/2; 436/2; 492/3; 657/3; 759/1; 820/1

##### Alignment Scores:

| Pred. No.:             | 1-21e-24 | Length:          |
|------------------------|----------|------------------|
| Score:                 | 451.50   | Matches: 158     |
| Percent Similarity:    | 39.38%   | Conservative: 96 |
| Best Local Similarity: | 24.50%   | Mismatches: 225  |
| Query Match:           | 14.52%   | Indels: 166      |
| DB:                    | 2        | Gaps: 20         |

US-09-270-437D-5 (1-1708) x T23837 (1-854)

|    |     |  |                                 |
|----|-----|--|---------------------------------|
| QY | 40  | CATCATGAGCTGAATGGCCACCATGGAGAACCATCCCTGAGGTCTCTCATACCC | 99                              |
|    |     |  |                                 |
| Db | 241 | HisHisGlnGln-HisProGlnMetMetGlnHisala---               | GlnGlnGlyTyrHisPr 259           |
|    |     |  |                                 |
| QY | 100 | CGATGAGCAGATAGCACACAG-----                             | GGACCTGAGATGGGCCCGCGGGGTTGG 150 |
|    |     |  |                                 |
| Db | 259 | chisGlnGlnAenGlnGlnHisGlnHisGlnHisGlnSerHisGlnSerGln   | 279                             |

|     |   |     |   |
|-----|---|-----|---|
| 151 | CTCTCGGGGTACGCCCGCCAGGGGCTCACTGTGTGGCAGCGGGGGCCCCCAGCAAGCAGCA | 211 | CTCTCGGGGTACGCCCGCCAGGGGCTCACTGTGTGGCAGCGGGGGCCCCCAGCAAGCAGCA |
| 152 |   | 212 |   |
| 153 |   | 213 |   |
| 154 |   | 214 |   |
| 155 |   | 215 |   |
| 156 |   | 216 |   |
| 157 |   | 217 |   |
| 158 |   | 218 |   |
| 159 |   | 219 |   |
| 160 |   | 220 |   |
| 161 |   | 221 |   |
| 162 |   | 222 |   |
| 163 |   | 223 |   |
| 164 |   | 224 |   |
| 165 |   | 225 |   |
| 166 |   | 226 |   |
| 167 |   | 227 |   |
| 168 |   | 228 |   |
| 169 |   | 229 |   |
| 170 |   | 230 |   |
| 171 |   | 231 |   |
| 172 |   | 232 |   |
| 173 |   | 233 |   |
| 174 |   | 234 |   |
| 175 |   | 235 |   |
| 176 |   | 236 |   |
| 177 |   | 237 |   |
| 178 |   | 238 |   |
| 179 |   | 239 |   |
| 180 |   | 240 |   |
| 181 |   | 241 |   |
| 182 |   | 242 |   |
| 183 |   | 243 |   |
| 184 |   | 244 |   |
| 185 |   | 245 |   |
| 186 |   | 246 |   |
| 187 |   | 247 |   |
| 188 |   | 248 |   |
| 189 |   | 249 |   |
| 190 |   | 250 |   |
| 191 |   | 251 |   |
| 192 |   | 252 |   |
| 193 |   | 253 |   |
| 194 |   | 254 |   |
| 195 |   | 255 |   |
| 196 |   | 256 |   |
| 197 |   | 257 |   |
| 198 |   | 258 |   |
| 199 |   | 259 |   |
| 200 |   | 260 |   |
| 201 |   | 261 |   |
| 202 |   | 262 |   |
| 203 |   | 263 |   |
| 204 |   | 264 |   |
| 205 |   | 265 |   |
| 206 |   | 266 |   |
| 207 |   | 267 |   |
| 208 |   | 268 |   |
| 209 |   | 269 |   |
| 210 |   | 270 |   |
| 211 |   | 271 |   |
| 212 |   | 272 |   |
| 213 |   | 273 |   |
| 214 |   | 274 |   |
| 215 |   | 275 |   |
| 216 |   | 276 |   |
| 217 |   | 277 |   |
| 218 |   | 278 |   |
| 219 |   | 279 |   |
| 220 |   | 280 |   |
| 221 |   | 281 |   |
| 222 |   | 282 |   |
| 223 |   | 283 |   |
| 224 |   | 284 |   |
| 225 |   | 285 |   |
| 226 |   | 286 |   |
| 227 |   | 287 |   |
| 228 |   | 288 |   |
| 229 |   | 289 |   |
| 230 |   | 290 |   |
| 231 |   | 291 |   |
| 232 |   | 292 |   |
| 233 |   | 293 |   |
| 234 |   | 294 |   |
| 235 |   | 295 |   |
| 236 |   | 296 |   |
| 237 |   | 297 |   |
| 238 |   | 298 |   |
| 239 |   | 299 |   |
| 240 |   | 300 |   |
| 241 |   | 301 |   |
| 242 |   | 302 |   |
| 243 |   | 303 |   |
| 244 |   | 304 |   |
| 245 |   | 305 |   |
| 246 |   | 306 |   |
| 247 |   | 307 |   |
| 248 |   | 308 |   |
| 249 |   | 309 |   |
| 250 |   | 310 |   |
| 251 |   | 311 |   |
| 252 |   | 312 |   |
| 253 |   | 313 |   |
| 254 |   | 314 |   |
| 255 |   | 315 |   |
| 256 |   | 316 |   |
| 257 |   | 317 |   |
| 258 |   | 318 |   |
| 259 |   | 319 |   |
| 260 |   | 320 |   |
| 261 |   | 321 |   |
| 262 |   | 322 |   |

## RESULT 2

D96554

hypothetical protein F19C24.19 [imported] - Arabidopsis thaliana

C/Species: *Arabidopsis thaliana* (mouse-ear cress)

C:\Date: 02-Mar-2001 #sequence revision 02-Mar-2001 #text change 31-Mar-2001

C:Accession: D96554

Reese, D. J.; Ecker, J.R.; Palm, C.J.; Pederspiel, N.A.; Kaul, S.; White, O.; Alonso, R.; Theologis, A.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Cressy, T.H.; Dewar, K.; Chin, C.W.; Hughes, B.; Huizar, L.; Janssen, N.F.; Janssen, B.F.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: D96554  
A;Status: preliminary

A; Status: preliminary  
A: Molecule type: DNA

A:Molecule type: DNA  
A:Residues: 1-621 <STO>

A;Residues: I-62I <SIO>  
A:Cross-references: GB:

A/CROSS-REFERENCES: 95:AE003173; NID:91107470Z; FIDN:FA023033.T; 93FDD:GN007471



C;Genetics:  
 A;Gene: F19C24.19  
 A;Map position: 1

Alignment Scores:  
 Pred. No.: 3,65e-11 Length: 621  
 Score: 265.00 Matches: 117  
 Percent Similarity: 38.66% Conservative: 74  
 Best Local Similarity: 23.68% Mismatches: 184  
 Query Match: 8.52% Indels: 119  
 DB: 2 Gaps: 19

US-09-270-437D-5 (1-1708) x D96554 (1-621)

QY 221 ATCCCCCTCGGTCCTGGTCCGCCACCCAGTATGTGGTCCCATTTGSCAAGAGGGG 280  
 Db 19 ValHisPheArgLeuLeuLeuCysProAlaThrArgThrGlyAlaLeuLeuGlyGly 38  
 QY 281 GCACCATCCGCAACATCACAACAGACCCCATCCCAAGATA----- 322  
 Db 39 SerValileArgHisLeuGlnSerValThrGlySerLysileArgValileAspAspIle 58  
 QY 322 ----- 322  
 Db 59 ProValProSerGluGluArgValValLeuLeuLeuAlaProSerGlyLysLysAsp 78  
 QY 323 -----GACGTGCATAGGAAGAGACCGCAGGTGCACGTGAA----- 358  
 Db 79 GluSerAsnValCysAspSerGluAsnProGlySerGluGluProLysGlnGluLysGly 98  
 QY 359 -----AAGCCATCAGTGTGCATCCACCCCTGAGGGCTGCTCCGCTTGTAA 409  
 Db 99 SerGluCysAlaGlyThrSerGlyGlyAspAspGluGluAlaProSerSerAlaGlnMet 118  
 QY 410 ATGATCTTGAGATTATGCAATAA-----GAGCTAAGGACACCAAAACGGCTGAC--- 460  
 Db 119 AlaLeuLeuArgValPheGluArgGluValPheGlyAspAspAlaAlaThrValAspGly 138  
 QY 461 -----GAGTTCCCTCGAAGTCTCTGCGCCCATATAACTTT 496  
 Db 139 AspGluLeuAspLysGlyGluSerGluGlyLeuCysArgMetileValArgGlyAsnGln 158  
 QY 497 GTAGGGCTCTCATTTGCAAGAGGACGAACTGAAGAAGGTAGACGAAGATACCGAG 556  
 Db 159 ValAspTyrLeuMetSerLysGlyGlyMetileGlnLysileArgGluAspSerGly 178  
 QY 557 ACAAATACCATCTCTCTGTTGCAA-----GACCTT 589  
 Db 179 AlaileValArgileSerSerThrAspGlnleProProCysAlaPheProGlyAspVal 198  
 QY 590 ACCCTTTACACCTCGAGGACCATCATCTGTGAGGGGGCC-----ATCGAGAT 640  
 Db 199 ValileGlnMetAsnGlyLysPheSerSerValLysLysAlaLeuLeuValThrAsn 218  
 QY 641 TGT-----TGACGGGCGGAGCAAGAAATA 664  
 Db 219 CysLeuGlnGluSerGlyAlaProProThrTrpAspGluCysProPheProGln----- 236  
 QY 665 ATGAAGAAAGTTCCGGAGGCTATGAGAATGATGCTGCCATGAGCTCTACCTGATC 724  
 Db 237 -----ProGlyTyrProProGluTyrHisSerMetGluTyrHis----- 249  
 QY 725 CTGGCTGAACCTGGCTGTGTAGGTCTTTCCAGCTTCATCCAGCGCAGTCCCGCG 784  
 Db 250 -----ProGlnTrpAspHis-----ProPro 256  
 QY 785 CCTCCACGACGTTACTGGGGCTGCTCCCTATAGCTCCTTTATGCAAGGCTCCCGAGCAG 844  
 Db 257 ProAsnProMetProGluAspValGlyProPheAsnArgProVal---ValGluGluGlu 275  
 QY 845 GAGATGTGCAAGTGTATTATCCCGCCAGGAGTGGGCCCATCATCCGACAGAGGGG 904  
 Db 276 ValAlaPheArgLeuLeuCysProAlaAspLysValGlySerLeuileGlyLysGly 295

QY 905 CAGCACATCAACACAGCTCTCCCGGTTTCCAGCGCTCCATCAAGATTGACCCACCGAA 964  
 Db 296 AlaValValArgAlaLeuGlnAsnGluSerGlyAlaSerIleLysValSerAspPro--- 314  
 QY 965 ACACCTGACTCCAAAGTTCGTATGTTATCATCTGACCGCCAGAGGCCCAATTCAG 1024  
 Db 315 ThrHisAspSerGluGluArgIleValIleSerAlaArgGluAsnLeuGluArgArg 334  
 QY 1025 -----GTCACGGGA-----AGAATCTATGCAAACTCAAGAGGAGGAGAACATTC 1066  
 Db 335 HisSerLeuAlaGlnAspGlyValMetArgValHisAsnArgIleValGluIleGlyPhe 354  
 QY 1067 TTTGTCTCCAAAGGAGGAAGTGAAGCTGCAGAGACCCACATACGTGTGCCAGCATCAGACGT 1126  
 Db 355 -----GluProSerAlaAlaValAlaArgLeuValHisSerProTyrIle 371  
 QY 1127 GCCCGGTTCATTCGCAAGGTGGAACGGTGAACGAGTTGCAGAAATTCAGCGCAGCT 1186  
 Db 372 GlyArgLeuLeuGlyLysGlyGlyHisLeuileSerGluMetArgAlaThrGlyAla 391  
 QY 1187 GAGGTG---GTAGTACCAAGAGACCCAGACCCCTGATGAGAACGACGAG-----GTCTATC 1237  
 Db 392 SerIleArgValPheAlaLysAspGlnAlaThrLysTyrGluSerGlnHisAspGluIle 411  
 QY 1238 GTGAAATCATCGACATTTCTATCCAGTCAGATGGCTCAACGGAAGATCCGAGACATC 1297  
 Db 412 ValGlnValIleGlyAsnLeu-----LysThrValGlnAspAlaPheGlnIle 428  
 QY 1298 CTGGCCCGAGTTAAGCAGCAGCATCAGAAGGACAGAGTAACGAGGCCCGCCAGCGAGG 1357  
 Db 429 LeuCysArgLeuArgGluAlaMetPheProGly-ArgLeuProPheGlnGlyMetGlyG 448  
 QY 1358 AAGTGACACGCCCTC-----CCTGTCTCTTNGAGTCCAGGACA 1396  
 Db 448 YProProProProPheMetGlyProTyrProGluProProProPheGlyProArgG 468  
 QY 1397 ACAACGGGCGAGAAATCGAGAGTGTCTCTCCCGCAGGC 1436  
 Db 466 nTyrProAlaSerProAspArgTyrHisSerProValGly 481

## RESULT 3

T41600  
 probable pre-mRNA splicing factor - fission yeast (Schizosaccharomyces pombe)  
 C;Species: Schizosaccharomyces pombe  
 C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
 C;Accession: T41600  
 R;lyne, M.; Rajandream, M.A.; Barrell, B.G.; Bothe, G.; Pohl, T.  
 submitted to the EMBL Data Library, October 1998  
 A;Reference number: Z22003  
 A;Accession: T41600  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-398 <LYN>  
 A;Cross-references: EMBL:AL031825; PIDN:CAA21234.1; GSPDB:GN00068; SPDB:SPCC757.09C  
 A;Experimental source: strain 972h-; cosmid c757  
 C;Genetics:  
 A;Gene: SPDB:SPCC757.09C  
 A;Map position: 3

## Alignment Scores:

Pred. No.: 1,34e-10 Length: 398  
 Score: 256.50 Matches: 93  
 Percent Similarity: 44.74% Conservative: 60  
 Best Local Similarity: 27.13% Mismatches: 130  
 Query Match: 8.25% Indels: 59  
 DB: 2 Gaps: 14

US-09-270-437D-5 (1-1708) x T41600 (1-398)

QY 161 CAGGCCCGC-----CAGGGCTCAGCTGTGGCAGCGGGGCCCA----- 199  
 Db 68 GlnProGluProThrSerGlnValProPheSerAlaLysProProMetAspAspAla 87



Db 228 HisLeu-----LeuLeuSerSerSerSerSer 237  
 Qy 776 GTCCGCGCGCTCCACAGCGTTACTGGGCTGCTCCCTATAGCTCCTTTATGACG--- 832  
 Db 238 MetHisGlnProGlyAlaMetLeuMetSerAlaAlaLeuThrSerHisArgAsnTyr 257  
 Qy 833 -----GTCCTCGGACGAGATGGTCAGGTGTTATC-----CCGCGC 871  
 Db 258 AlaValArgArgAspIleAlaAspAlaArgGluPheCysValCysPheIleCysProAla 277  
 Qy 872 CAGCGAGTGGCGCGCATCATCGGCAAGAGGGCGCAGCACATCAACAGCTCTCCCGGTTT 931  
 Db 278 GluAsnValGlyGlyValIleGlyGlyGlyPheIleAsnGlnIleArgGlnGlu 297  
 Qy 932 GCACGCGCTCCATCAAGATTGACACCCCGCAACACCTCACTCCAAAGTTCGTATGGTT 991  
 Db 298 ThrGlyAlaThrIleArgValAsnThrSerGluThrAspAspAsp-----CysIle 315  
 Qy 992 ATCATCACTGGACCGCCAGAG-----GCCCAATTC 1021  
 Db 316 IlePheIleSerSerLysGluPheTyrGluAspGlnSerProAlaValAsnAlaIle 335  
 Qy 1022 AAGGCTCAGGAGAAATCTATGGCAAACTC---AAGGAGGAGAACTTCTTTGTCCTCCAAAG 1078  
 Db 336 ArgLeuGlnGlnArgCysSerGluLysValGlyLysAspAlaAsn----- 350  
 Qy 1079 GAGGAAGTGAAGCTGGAGACCCATACAGTGTGCCAGCATCAGCAGCTGCGCGGTCATT 1138  
 Db 351 ---AspLeuAlaIleSerThrArgLeuLeuValSerSerSerGlnIleGlyCysLeuIle 369  
 Qy 1139 GGCAAGTGGAAAAACGGTGAACAGTTCGCAAAATTTGACGGCAGCTGAGGTG---GTA 1195  
 Db 370 GlyLysGlyGlyAlaValIleSerGluMetArgSerValThrArgAlaAsnIleArgIle 389  
 Qy 1196 GTACCAAGACACGACCCCTGAT-----GAGAACGACAGCTCATCGTGAATAATCATC 1249  
 Db 390 LeuGlnLysGluAspValProLysIleAlaArgGluAspGluGluMetValGlnIleThr 409  
 Qy 1250 GGACATTCTATGCGCAGTCAGATGGCTCAACGAAAGATCCGAGACATCTCGGCCAGGTT 1309  
 Db 410 Gly-----SerProAspAlaAlaMetLysAlaLeuThrGlnValIleLeuArgLeu 426  
 Qy 1310 AAG 1312  
 Db 427 Arg 427

## RESULT 5

hypothetical protein ZK418.9 - Caenorhabditis elegans  
 C;Species: Caenorhabditis elegans  
 C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C;Accession: T27855  
 R;Fulton, L.  
 submitted to the EMBL Data Library, April 1994  
 A;Description: The sequence of C. elegans cosmid ZK418.  
 A;Reference number: Z20430  
 A;Accession: T27855  
 A;Status: preliminary;  
 A;Molecule type: DNA  
 A;Residues: 1-768 <FULL>  
 A;Cross-references: EMBL:U00047; PIDN:AA50693.1; CESP:ZK418.9  
 A;Experimental source: strain Bristol N2  
 C;Genetics:  
 A;Gene: CESP:ZK418.9  
 A;Introns: 166/1; 221/3; 243/1; 279/1; 387/3; 490/1; 567/3; 727/2  
 Alignment Scores:  
 Pred. No.: 6.59e-10 Length: 768  
 Score: 248.00 Matches: 131  
 Percent Similarity: 37.59% Conservative: 88  
 Best Local Similarity: 22.66% Mismatches: 225  
 Query Match: 7.97% Indels: 134  
 DB: 2 Gaps: 22

US-09-270-437D-5 (1-1708) x T27855 (1-768)  
 Qy 131 GGGCGCCGAGGGGGTGGTGGTCTCGGGTTCAGCCCGCCAGGGGTACCTGTGGCAGCG 190  
 Db 224 GlyAlaAspGlyGlyIleGlyThrLeuLysArgProLeuAsp---SerGluIleLeuAsp 242  
 Qy 191 GGG-----GCCCGAGCCCAAGCAGCAGCAA----- 214  
 Db 243 GlyAspLeuIleProThrLysLysSerSerGluValGlyAspLeuAsnMetGlyAspSer 262  
 Qy 215 -----GTGGACATCCCTTCGGCTCCTGGTCCCGCCAGTATGTGGTGGTCCATT 265  
 Db 263 AspLysIleThrAspIle-----TyrProValProGluLysValIleGlyLeuVal 279  
 Qy 266 ATTGCAAGGAGGGGCCACCATCCGCAACATCAACAAACAGACCCAGTCCAAAGTAGAC 325  
 Db 280 IleGlyLysGlyGlySerGluIleArgLeuIleGlnThrSerGlyCysArgValGln 299  
 Qy 326 GTGCATAGGAGGAGAACGCGAGTGCAGCTGAAAGGCCATCAGTGTGCATCCACCCCT 385  
 Db 300 MetAspProAspHisGlnSerValAsnGlyPheArgAsnCysThrIleGluGlyProPro 319  
 Qy 386 GAGGCTCTCTCTCCGCTTGTAAAGATGATCTTGAGAGATTATGCATAAAGAG----- 436  
 Db 320 AspGlnValAlaValAlaArgGlnMetIleThrGlnValIleAsnArgAsnGlnThrGly 339  
 Qy 437 GCTAAGGACACCAAAACGCGTACGAGGTTCCTGGAAGATCCTGGCCCAATAAATCTTT 496  
 Db 340 AlaGlnProGlyAlaAlaProGlyGluValThrGluGluMetLeuIleProAlaAspLys 359  
 Qy 497 GTAGGGCTCTCATGGCAAGGAGGAGCGAACCTGAAGAAGGTAGACAAAGATACCCGAG 556  
 Db 360 IleGlyLeuValIleGlyLysGlyGlyGluThrIleArgIleValGlnGluGlnSerGly 379  
 Qy 557 ACAAAATCAACATCTCTCTGTGTCAGACCTTACCTTTAC-----AACCTT 604  
 Db 380 LeuArg---AsnCysAsnValValGlnGluThrThrAlaThrGlyGlnProLysPro 398  
 Qy 605 GAGAGACCATCACTGTGAGGGGCCCATCGAGAATTGTCAGGGCCGAGCAGGAATA 664  
 Db 399 LeuArgMetIleGlySerProAlaAlaIleGluThrAlaLysAlaLeuValHisAsnIle 418  
 Qy 665 ATGAAGAAAGTTTCGGGAGCGCTATGAGAATGATGTGGCTGCCATGAGCTCTCACCTGATC 724  
 Db 419 MetAsnAsnThrGln----- 423  
 Qy 725 CCGTCCCTGAACCTGGCTGTAGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCG 784  
 Db 424 -----GlyAsnAlaProLeuLeuGlnArgAla---ProHis 434  
 Qy 785 CCGCCAGCAGCGCTTACTGGGCTCTCCCTATAGCTCTTTATGACGGCTCCCGAGCAG 844  
 Db 435 GlnProSerGlyGlnPheGlyGlyGlyTyrGlyAlaGlnGluAlaGlnAlaLysGly--- 453  
 Qy 845 GAGATGGTGCAGGTGTTTATCCCGCCCGCAGCAGTGGCGCCATCATCGCAAGAGGGG 904  
 Db 454 -----GluValIleValProArgLeuSerAlaGlyMetIleIleGlyLysGly 470  
 Qy 905 CAGCACATCAACACAGCTCTCCGGTTTCCAGCGCTCCATCAAGATTGACCCACCCGAA 964  
 Db 471 GluMetIleLysArgLeuAlaGluThrGlyThrLysIleGlnPheLysProAspThr 490  
 Qy 965 ACACCTGATCCAAAGTTCGTATGTTATCATCTGACCCCGCAGCCGACGAGCCCAATTCAG 1024  
 Db 491 AsnProAsnSerGluAspArgIleAlaValIleMetGlyThrArgAspGlnIleTyrArg 510  
 Qy 1025 GCTCAGGGAAGAACTATCGCAAACTCAAGAG-----GAGAATCTTTGGTCCCCAAG 1078  
 Db 511 AlaThrGluArgIleThrGluIleValAsnArgAlaIleLysAsnAsnGlyAlaProGln 530  
 Qy 1079 GAGGAAGTGAAGCTGGAGACC-----CACATAGTGTGCCA 1114  
 Db ---

[illegible]

## RESULT 6

myc far upstream element-binding protein - human  
 N;Alternate names: FUSE-binding protein  
 C;Species: Homo sapiens (man)  
 C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 05-Nov-1999  
 C;Accession: A53184  
 R;Duncan, R.; Bazar, L.; Michelotti, G.; Tomonaga, T.; Krutzsch, H.; Avigan, M.; Levens,  
 Genes Dev. 8, 465-480, 1994  
 A;Title: A sequence-specific, single-strand binding protein activates the far upstream  
 A;Reference number: A53184; PMID:94170991; PMID:8125259  
 A;Accession: A53184  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-644 <DUN>  
 A;Cross-references: GB:005040; NID:G460151; PIDN:AAA17976.1; PID:G460152  
 C;Keywords: DNA binding

QY 1019 TTCAAGGCTCAG----- 1030  
 Db 334 HisAlaAlaGluLeuIleThrAspLeuLeuArgSerValGlnAlaGlyAsnProGlyGly 353  
 QY 1031 -----GGAAGAATCTATGGCAAACTCAAGGAGGAGAACTTCTTGCTCCCAAG 1078  
 Db 354 ProGlyProGlyArgGlyArgGlyArgGlyGlnGlyAsnTrpAsnMetGlyProPro 373  
 QY 1079 GAGAAAGTGAAGTGGAGCCCATACGTGTGCCAGATCAGACGTGCGCGGFCAT 1138  
 Db 374 GlyGlyLeuGln---GluPheAsnPheIleValProThrGlyLysThrGlyLeuIleIle 392  
 QY 1139 GCAAAAGGTGAAAACGGTACAGAGTTCGCAAAATTCAGCGACCTCAGGTGGTAGTA 1198  
 Db 393 GlyLysGlyGluThrIleLysSerIleSerGlnGlnSerGlyAlaArgIleGluLeu 412  
 QY 1199 CCAAGAGACAGACCCCTGATGAAACAGCACCGAGTC---ATCGTGAATATCATCGACAT 1255  
 Db 413 GlnArgAsnProProProAsnAlaAspProAsnMetLysLeuPheThrIleArgGlyThr 432  
 QY 1256 -----TTCTATGCCAGTCAGATGGCTCAACGGAGAT-----CCGAGACAT 1296  
 Db 433 ProGlnGlnIleAspTyAlaArgGlnLeuIleGluGluLysIleGlyGlyProValAsn 452  
 QY 1297 CCT---GGCCCAAGGTAAAGCAGCAGCATCAGAAGGGACACAGTAACAGGCCCGACGACG 1353  
 Db 453 ProLeuGlyProProValProHisGlyProHisGlyVal-----ProGlyPro-HisG 470  
 QY 1354 GA-----GGAGTGCACAGCCCTC-----CTGTCCCTTNGA 1386  
 Db 470 LysProGlyProProGlyProGlyThrProMetGlyProTyAsnProAlaProTyAla 490  
 QY 1387 GTCCAGGA----- 1394  
 Db 490 snProGlyProProGlyProAlaProHisGlyProProAlaProTyAlaProGlnGlyT 510  
 QY 1395 -----CAACAACGGCAGAAATTCGAGAGTGTCTCTCCCGG 1431  
 Db 510 rpGlyAsnAlaTyProHisTrpGlnGlnAlaProProAsp-----ProA 526  
 QY 1432 CAGGCTGAGAAATGAGTGGAAATCCGGGACACNTGGCGGGCTGTAGATCAGGTGGC 1491  
 Db 526 lAlysAlaGlyThrAspProAsnSerAla-AlaAlaAlaAlaTy-----TyAla 542  
 QY 1492 CACTTGATTCAGAAAGATGTTCCAGTGAGGAACCTCTATNTCAGCCCAACACCCAC 1551  
 Db 543 HisTyTyGlnGlnGlnAla-----GlnProProAlaAla 555  
 QY 1552 CCAATTGGC---CCAACACTGTNTGCCCTCGGGGTGTGAGAAATNTAGCGCAAGGCAC 1608  
 Db 556 ProAlaGlyAlaProThrThrThrGlnThrAsnGlyGlnGlyAsp----- 570  
 QY 1609 TTTTAACGTGGATTGTTTAAAGAGCTCTCCAGGCCCAACCAAGAGGGTGGATCACAC 1668  
 Db 571 -----GlnGlnAsnProAlaProAlaGlyGlnValAspTyThr 583  
 QY 1669 TCAGTCGGGAGAAATAATAATTTCTTCAGT 1701  
 Db 584 LysAlaTrpGluGluTyTyLysLysMetGlyGlnAlaValProAlaProThrGly 602

## RESULT 7

T25832  
 hypothetical protein M01A10.1 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C:Accession: T25832  
 R:Scheet, P.  
 submitted to the EMBL Data Library, February 1997  
 A:Description: The sequence of C. elegans cosmid M01A10.  
 A:Reference number: Z20094  
 A:Accession: T25832  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA

A:Residues: 1-680 <SCH>  
 A:Cross-references: ENBL:U08174; PIDN:AAB42272.1; GSPDB:GN00019; CESP:M01A10.1  
 A:Experimental source: strain Bristol N2; clone M01A10  
 C:Genetics:  
 A:Gene: CESP:M01A10.1  
 A:Map position: 1  
 A:Introns: 63/2; 97/3; 121/1; 160/1; 269/3; 411/3; 512/1; 649/2

Alignment Scores:  
 Pred. No.: 2,23e-09 Length: 680  
 Score: 240.50 Matches: 82  
 Percent Similarity: 41.21% Conservative: 82  
 Best Local Similarity: 20.60% Mismatches: 159  
 Query Match: 7.73% Indels: 75  
 DB: 2 Gaps: 12

US-09-270-437D-5 (1-1708) x T25832 (1-680)

QY 104 GAGCAGATGACACAGGACCTGAGATGGGCGCCGAGGGGCTTTGGCTCTCGGGGTGAG 163  
 Db 97 LysHisIleSerSerSerThrAspThrAlaArgLysArgAspPheAspGluArgSer--- 115  
 QY 164 CCGCGCCAGGCTCACCT-----GTGGCAGCGGGGGCCCGCAGCCAGACGACGACGACGTG 217  
 Db 116 -----GluGlySerAspGluTyArgGluTyAlaProProCysLysLeuThrLysGly 133  
 QY 218 GACATCCCTTCGG-----CTCCTGGTGCC 244  
 Db 134 AspIleAspTyArgValAspThrSerThrThrValIleLysAlaSerValSerIlePro 153  
 QY 245 ACCCAGTATGTGGTGCATTTATTGGCAAGGAGGGGCCACCATCCGCAACATCACAAAA 304  
 Db 154 GluGluSerValGlyLeuValIleGlyArgAsnGlyValGluIleGlnAlaIleSerGln 173  
 QY 305 CAGACCCAGTCCAGATAGACGTGCATAGGAGGAGACGCGAGTGCAGCTGAAAAAGCC 364  
 Db 174 LysSerGlyCysArgValGlnIle---ValAlaGluProSerThrThrGlyTyArgSer 192  
 QY 365 ATCAGTGTGACTCCACCCCTGAGGGCTGCTCCCTGTAAGATGATCTTGAGATT 424  
 Db 193 ValAspIleTyGlyIleSerGluAsnIleGluValAlaLysLysLeuIleAsnGluVal 212  
 QY 425 ATGCATAAAGAGGTAAAG----- 442  
 Db 213 ValAlaArgGlyArgLysLeuSerGlnGluProLeuProCysSerValProGlnPheGln 232  
 QY 443 -----GACACCAAAACGCTGACGAGTTCCTCTGAAGATCCTCGGCCCAATAAATAC 493  
 Db 233 ProIleProAlaValSerAsnSerSerLysValThrIleIleProIleProAlaAsn 252  
 QY 494 TTTGTAGGCGCTCTCATTTGGCAAGAGGACGCGAACCTGAAGAAGGTAGACGAAGATACC 553  
 Db 253 LysCysGlyAlaIleIleGlyLysLysGlyGluGlnMetArgLysLeuArgSerTrpThr 272  
 QY 554 GNACAAAAATCACCATCTCTCGTTGCAAGACCTTACCTTTACACCCCTGAGAGGACC 613  
 Db 273 AsnCysAspPheIleLeu-----IleGlnGluAsnAsnIleAlaAspSerValLysPro 290  
 QY 614 ATCACTGTGAAGGGGCCCATCGAATTTGTCAGGGCCGACGAGAAATATTAAGAAAA 673  
 Db 291 LeuGlnIleThrGlyGln-----ProLysGlu 299  
 QY 674 GTTCGGGAGGCTATGAGAATGATGTGCTGCATGAGCTCTACCTGATCCTGGGCTG 733  
 Db 300 ValGluHisAla-----LysAlaLeuValAlaAspIleLeuAspGlyPhe 314  
 QY 734 AACCTGGCTGCTGTAGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCCGCTCCACGC 793  
 Db 315 AspGluCys-----ProProAla 320  
 QY 794 AGCGTTACTGGGCTGCTCCCTATAGTCTCTTTATGACGGTCCGACGACGAGATGTGTG 853  
 Db 321 GlyMetAlaGlyAsnSerProValAlaAlaMet-----SerLeu 333



```

386 -----AlaGlnGlyGlyTyrAspTyrTyrGlyGlnGlnGlnSerGlnGlnProSerSer 403
QY 997 CACTGGACGCCAGAGCCCAATTAAGGCTCAGGGAAGAACTATGCGCAAACTCAAGGA 1056
Db 404 GlyGlySerSerAlaProProThrAspThrThrGlyTyrAsnTyrTyrGlnHisAlaSer 423
QY 1057 GGAAAGACTTCTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACCCACATACGTGTGCCAGC 1116
Db 424 GlyTyrGlyGlnAlaGlyGlnGlyTyrGlnGln-----AspGlyTyrGlyAlaTyrAsn 441
QY 1117 ATCAGCAGCTGGCGGGTCTATGGCAAGGTGGAAGGTGGAACCGTGAACGACTTCAGAAATT 1176
Db 442 AlaSerGlnGlnSerGlyTyrGlyGln-----AlaAlaGlyTyr 454
QY 1177 GAGCGCAGCTGAGCT---GCTAGTACCAAGACAGACCCCTGATGAGAAAGACCGAGCT 1233
Db 455 AspGlnGlnGlyTyrGlySerThrThrAsnPro-SerGlnGlnGluAsp-----471
QY 1234 CATCGTGAAATCATCGGACATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCGGAGA 1293
Db 472 -----AlaSerGlnAlaAlaPro-----477
QY 1294 CATCCTGGCCAGGTTAAGCAGCAGCATCAGAGGAGCAGAGTAACACGAGCCAGGCAGC 1353
Db 478 -----ProSerSerAlaGlnSerGlyGlnAlaGlyTyrGlyThrThrGln 492
QY 1354 GAGGAAGTGACACGCCCTCCCTGCTCCCTNGAGTCCAGGACCAACACGGCGCAAAATCG 1413
Db 492 yGlnGlnProAlaGlnGlySerThrGlyGlnAlaGlyTyrGlyAlaProProThrSe 512
QY 1414 AGAGTGTCTCTCCCGGAGGCTGAGATGATGAGTGGGATCGGACACNTGGGCGCGG 1473
Db 512 rGlnAlaGlyTyrSerSerGlnProAlaAlaTyrAsnSerGlyTyrGlyAlaProPr 532
QY 1474 CTGTAGATCAGGTTTGCCCACT 1495
Db 532 oProAlaSerLysProProThr 539

RESULT 9
T49962
hypothetical protein F8M21.160 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C;Accession: T49962
R;Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Le
submitted to the Protein Sequence Database, April 2000
A;Reference number: 224493
A;Accession: T49962
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-568 <BR>
A;Cross-references: EMBL:AL353993; GSPDB:GN00063; ATSP:F8M21.160
A;Experimental source: cultivar Columbia; BAC clone F8M21
C;Genetics:
A;Gene: ATSP:F8M21.160
A;Map position: 5
A;Introns: 200/3; 337/3; 544/3

Alignment Scores:
Pred. No.: 5,78e-09 Length: 568
Score: 234.50 Matches: 92
Percent Similarity: 42.47% Conservative: 97
Best Local Similarity: 20.67% Mismatches: 175
Query Match: 7.54% Indels: 81
DB: 2 Gaps: 16

US-09-270-437D-5 (1-1708) x T49962 (1-568)
QY 98 CCCGATCAGCAGATAGCAGAGGACCTGAGAATGGGCGCGCAGAGGGGCTTTGGCTCTCGG 157
Db 3 ProAspHisArgMetSerProAspHisArgAspSerHisArg-----LysArg 18
```

```

158 GGTGAGCCCGCCAGGGCTCACCTGTGGCAGCGGGGGCCCCAGCAACAGCAGCAG----- 211
Db 19 SerArgProGlnSerAspTyrAspAspAsnGlyGlySerLysArgAlaGlyTyrArgGlyAsp 38
QY 212 -----CAAGTGGACATCCCCCTTCGGCTCCGTGGTCCGCCACC 247
Db 39 AspArgAspSerLeuValIleAspArgAspAspThrValPheArgTyrLeuCysProVal 58
QY 248 CAGTATGTGGTGCCATTTATTGGCAAGGAGGGGCCACCATCCGCAACATCACAAAAACAG 307
Db 59 LysLysIleGlySerValIleGlyArgGlyAspIleValLysGlnLeuArgAsnAsp 78
QY 308 ACCAGTCCAAAGATAGACGTGCATAGGAAGAGAACCGAGTGCAGCTGAAAAAGCCATC 367
Db 79 ThrArgSerLysIleArgIle---GlyGluAlaIleProGlyCysAspGluArgValIle 97
QY 368 AGTGTGCACTCCACCCCT-----GAGGGGTGCTCTCTCC 400
Db 98 ThrIleTyrSerProSerAspGluThrAsnAlaPheGlyAspGlyGlyLysValLeuSer 117
QY 401 GCTTGTAAAGATGATTTGGAGATTATGATAAAGAG-----GCTAAGGACACCAAAACG 454
Db 118 ProAlaGlnAspAlaLeuPheArgIleHisAspArgValValAlaAspAspAlaArgSer 137
QY 455 GCTGAC-----GAGTTCCCTGAAGATCCTGGCCCATATAAATCATT 496
Db 138 GluAspSerProGluGlyGlyGlnValThrAlaLysLeuLeuValProSerAspGln 157
QY 497 GTAGGGCTCTCATTTGGCAAGAGGACCGAACCTGAAGAGGTAGAGCAACATCCCGAG 556
Db 158 IleGlyCysIleLeuGlyArgGlyGlnIleValGlnAsnIleArgSerGluThrGly 177
QY 557 ACAAAATCACCATCTCTCGTTCAGACCTTACCCTT-----TACAACTTGAGAGG 610
Db 178 AlaGlnIleArgIleValLysAspArgAsnMetProLeuCysAlaLeuAsnSerAspGlu 197
QY 611 ACCATCACTGTGAAGGGGCCATCGAATTTGTGCAGCGCCGAGCAAGAAATATGTAAG 670
Db 198 LeuIleGlnIleSerGlyGluValLeuIleValLysLysAlaLeuLeuGlnIleAlaSer 217
QY 671 AAAGTTCGGAGGCGCTATGAGAATGATGTGCTGCCATGAGCTCTCACTGATCCCTGGC 730
Db 218 ArgLeu-----HisGluAsn-----ProSerArgSerGlnAsnLeu----- 229
QY 731 CTGAACCTGGCTGTGTAGTCTTTTCCACGCTTCATCCAGCGAGTCCCGCGCTCC 790
Db 230 -----LeuSerSerSerGlyTyrProAlaGlySerLeuMetSerHisAlaGlyGly 247
QY 791 AGCAGCGTTACTGGGGCTGCTCCCTATAGCTCC-----823
Db 248 ProArgLeuValGlyLeuAlaProLeuMetGlySerTyrGlyArgAspAlaGlyAspTrp 267
QY 824 -----TTATGACAGCTCCCGAGCAGGAG-----ATGCTGCAG 856
Db 268 SerArgProLeuTyrGlnProProArgAsnAspProProAlaThrGluPhePheIleArg 287
QY 857 GTGTTTATCCCGCCAGGCGAGTGGCGCCCATCATCGCAAGAGGGGCGAGCATCAAA 916
Db 288 LeuValSerProValGluAsnIleAlaSerValIleGlyLysGlyGlyAlaLeuIleAsn 307
QY 917 CAGCTCTCCCGGTTGGCCAGCGCTCCATCAAGATTGCAACACCGAAACACCTGACTCC 976
Db 308 GlnLeuArgGlnGluThrArgAlaThrIleLysValAspSerSerArgThrGluGlyAsn 327
QY 977 AAAATTCTGTATGGTTATC-----ATCACTGGACCGCCA 1009
Db 328 AspCysLeuIleThrIleSerAlaArgGluValPheGluAspAlaTyrSerProThrIle 347
QY 1010 GAGGCCCAATTCAAGGCTCAGGGAAGATCTATGGCAAACTCAAGAGGAGAACTTTT 1069
Db 348 GluAlaValMetArgLeuGlnProLysCysSerAspLysValGluArgAspSerGlyLeu 367
QY 1070 GGTCCCAAGGAGAGTGAAGCTGAGACCCACATACCTGTGCCAGCATCAGCAGCTGGC 1129
```



Db 368 -----ValSerPheThrThrArgLeuLeuValProSerArgIleGly 382  
 Qy 1130 CGGTCATTCGCAAGGTGAAACCGTGAACAGTTGCGAATTTGACGGCAGCTGAG 1189  
 Db 383 CysileLeuGlyysGlyAlaIleThrGluMetArgMetThrLysAlaAsn 402  
 Qy 1190 GTG---GTAGTACCAAGACAGACCCCT-----CATGAGAACACAGAGTC 1234  
 Db 403 IleArgIleLeuGlyysGluAsnLeuProLysValAlaSerAspAspGluMetVal 422  
 Qy 1235 ATCGTGAAATCATC 1249  
 Db 423 GlnValAsnPheMet 427  
 RESULT 10  
 T19216  
 Hypothetical protein C12D8.1a - Caenorhabditis elegans  
 C/Species: Caenorhabditis elegans  
 C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 R/McMurray A  
 submitted to the EMBL Data Library, June 1996  
 A/Reference number: Z19092  
 A/Accession: T19216  
 A/Status: preliminary; translated from GB/EMBL/DDBJ  
 A/Molecule type: DNA  
 A/Residues: 1-589 <WIL>  
 A/Cross-references: EMBL:Z73969; PIDN:CAA98232.1; GSPDB:GN00023; CESP:C12D8.1a  
 A/Experimental source: Clone C12D8  
 C/Genetics:  
 A/Gene: CESP:C12D8.1a  
 A/Map position: 5  
 A/Introns: 7/3; 25/1; 159/3; 318/3; 513/2  
 Alignment Scores:  
 Pred. No.: 8,86e-09 Length: 589  
 Score: 232.00 Matches: 108  
 Percent Similarity: 37.15% Conservative: 106  
 Best Local Similarity: 18.75% Mismatches: 232  
 Query Match: 7.46% Indels: 130  
 DB: 2 Gaps: 18  
 US-09-270-437D-5 (1-1708) x T19216 (1-589)  
 Qy 197 CCAGCCAGCAGCAGCAGCAGTGGACATCCCCCTT----- 229  
 Db 28 ProAlaLysArgProAlaAspAspThrAspLeuAsnProPheMetAspAspAsnGluAla 47  
 Qy 230 -----CGGTCCTGTGTCGCCACCCAGTATGTGGTGCCATTATTGGCAAGAGGG 280  
 Db 48 ValAsnGluLysTyrProIleProGluSerAlaValGlyIleValIleGlyArgGlyGly 67  
 Qy 281 GCCACCATCCGCACATCAACAAACACAGCCAGTCCAGATAGAGCTGCATAGGAAGGAG 340  
 Db 68 SerGluIleGlnGlyIleGlnAlaLysAlaGlyCysArgValGlnMetSerProAspAla 87  
 Qy 341 AACCGAGGTGCAGCTGAAAAGCCATCAGTGTGCTCCACCTCCAGGGTGTCTCTCC 400  
 Db 88 AspProSerSerGlyValArgMetValThrLeuGluGlySerArgSerAsnValGluThr 107  
 Qy 401 GCTGTGAAGATGATCTGTGAGATTATGCATATAAGAGGCTTAAGACACCAAA----- 451  
 Db 108 AlaLysHisLeuIleAsnGluValValAlaArgSerGlnAsnProArgProGlnTyrGly 127  
 Qy 452 ---ACGGTGCAGAGTTCCCTGGAAGATCTGGGCCATAATPAATTTGTAGGGCGTCTC 508  
 Db 128 PheProArgAlaGlnThrThrIleAspIleAlaIleProProAsnArgCysGlyLeuIle 147  
 Qy 509 ATTGGCAAGAGGAGGAACTGGAAGAGGTAGAGAGATACCGAGACAAATATCACC 568  
 Db 148 IleGlyLysSerGlyAspThrIleArgGlnLeuGlnGluLysSerGlyCysLysMetIle 167

Qy 569 ATCTCCTCTGTTGCAAGAC---CTTACCTCTTTTCAACCTCTGAGAGGACCATCTACTGTGAAG 625  
 Db 168 Leu-----ValGlnAspAsnGlnSerValSerAspGlnSerLysProLeuArgIleThr 185  
 Qy 626 GGGCCATCGAGAATTGTTGAGGGCCGAGCAGGAAATTAATGAAGAAGTTTCGGAGGCC 685  
 Db 186 GlyAspProGlnLysIleGluLeuAlaLysGlnLeuValAlaGluIleLeuAsnSerGly 205  
 Qy 686 TATGAGATGATGTGGTGCATCAGCTCTCACTGATCCCTGCTGCTGCTGCTGCTGCTGCT 745  
 Db 206 GlyAspGlyAsnGlyGlySer-----GlyLeuGlnMetHisHis 218  
 Qy 746 GTAGTCTTTTCCAGCTTTCATCCAGCGCAGTCCCGCGCTCCACGACGCTTACTCTGGG 805  
 Db 219 AlaGly-----Gly 221  
 Qy 806 CTCTCTCCCTATAGCTCTTTTATGACGGCTCCGAGCAGGAGATGTTGTCAGGTGTTTATC 865  
 Db 222 GlyGlyGlyAlaSerAla-----ArgGlyGluValValVal 233  
 Qy 866 CCCGCCAGGAGGCGGCATCATCGCAAGAGGGCGAGCAGCACAACACAGCTCTCC 925  
 Db 234 ProArgSerValGlyIleIleIleGlyLysGlnGlyAspThrIleLysArgLeuAla 253  
 Qy 926 CGTTTCCAGCGCTCCATCAAGATTGACACCCGAAACACCTCACTCCCAAGTTCTGT 985  
 Db 254 MetGluThrGlyThrLysIleGlnPheLysProAspAspAspProSerThrProGluArg 273  
 Qy 986 ATGTTTATCATCTGACCGCCAGAGGCCCAATTCAGGCTCAGGAGAAAGATCTATGCG 1045  
 Db 274 CysAlaValIleMetGlyThrArgAspGlnIleTyrArgAlaThrGluArgIleThrGlu 293  
 Qy 1046 AACTCAAGGAGGAG-----AACTCTTCTTGTGTCACAGGAG 1081  
 Db 294 LeuValLysLysSerThrMetGlnGlnGlyGlyGlyAsnValAlaGlyAlaMetVal 313  
 Qy 1082 GAAGTGAAGTGGAGACC---CACATAGTGTGCCAGCATCATCAGAGTGGCCGGGTCAAT 1138  
 Db 314 SerAsnGluAlaSerThrPheTyrMetSerValProAlaAlaLysCysGlyLeuValIle 333  
 Qy 1139 GGCAGAGGTGGAAACCGGTGAACGAGTTCAGAAITTCAGCGCAGTGGGTGGTAGTA 1198  
 Db 334 GlyLysGlyGlyGluThrIleLysGlnIleAsnSerGluSerGlyAlaHisCysGluLeu 353  
 Qy 1199 CCAAGAGACGAGACCCCTGATGAGAACGACGACAGTTCATCTGT-----AAA 1243  
 Db 354 SerArgAspProThrGlyAsnAlaAspGluLysValPheValIleLysGlyGlyLysArg 373  
 Qy 1244 ATCATCCGACATTTCTATGCCAGTCAGTGGCTCAACGGAGAGATCCGAGCATCTCGGCC 1303  
 Db 374 AlaIleGluHis-----AlaLysHisLeuIleArgIleLysValGlyAspIleAlaPro 391  
 Qy 1304 CAGGTGAAGCAGCAGCATCAGAAAGGGACAGTAACGAGCCCGCCAGGACGAGGAAGTGA 1363  
 Db 392 AsnThrProPheArgAspAspSerAlaMetThrMetGlnThrGlnPhe----- 407  
 Qy 1364 CCAGCCCTCTCTGTCTCCCTTNGAGTCAGGACAAACCGGGCAGAAATCGAGAGTGTGCT 1423  
 Db 408 ---SerAlaProAlaGlnAsnAsnPheGlyGlyGlnGlnGlnTrpAsnProValAlaGln 426  
 Qy 1424 CTCGCCGCGAGGCTCAGAAATGAGTGGGAATCCGG----- 1459  
 Db 427 IleProAlaAlaAlaGlnAsnProTyrGlnValGlyGlyTrpGlnGlnAsnSerValTyr 446  
 Qy 1460 -----ACACNTGGCGCGGTGTAGATCAGGTTTG----- 1489  
 Db 447 AlaGlnGlnThrAlaAlaProAlaAlaAlaProTyrAlaAlaAlaGlyIleValGlnPro 466  
 Qy 1490 -----CCCACCTTGATTCGAAAGATGTTCCAGTGGAGAAC 1525  
 Db 467 GlnGlnGlnValAlaTyrGlnGlnProGlnValValGlnGlnVal---GlnProAlaThr 485  
 Qy 1526 CTGATCTNTCAGCCCAACACCCACC----- 1552



```

Db      486 ThrValAlaAlaThrThrThrProThrValAlaProValThrGlyGluGlnAspTyrSer 505
QY      1553 ---CAATTGGCCCAACACTGTNTGCCCTCCGGGTG----- 1585
Db      506 AlaGlnTrpMetGluTyrTyrLysSerIleGlyAlaHisAspLysAlaGluAlaValGlu 525
QY      1586 -----TCAGAAATNTAGCCCAAGGCACCTTTAAACGTGATTG 1624
Db      526 AlaGlnMetLysLysLysAlaGluAlaAlaAlaArgAlaValProGlyGlyLeu 545
QY      1625 TTTTAAAGAGCTCTCCAGGCCCAACCAAGAGGGTGCATCACACCTCAG 1672
Db      546 -----IleGlnGlnMetProMetGlyMetAlaMetProGln 557

RESULT 11
T19217
hypothetical protein C12D8.1b - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C/Accession: T19217
R/McMurray, A.
submitted to the EMBL Data Library, June 1996
A/Reference number: Z19032
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-611 <MIL>
A/Cross-references: EMBL:Z73969; PIDN:CAA98233.1; GSPDB:GN00023; CESP:C12D8.1b
A/Experimental source: clone C12D8
C/Genetics:
A/Gene: CESP:C12D8.1b
A/Map position: 5
A/Introns: 181/3; 340/3; 535/2

Alignment Scores:
Pred. No.: 8,95e-09 Length: 611
Score: 232.00 Matches: 108
Percent Similarity: 37.15% Conservative: 106
Best Local Similarity: 18.75% Mismatches: 232
Query Match: 7.46% Indels: 130
DB: 2 Gaps: 18

US-09-270-437D-5 (1-1708) x T19217 (1-611)
QY      197 CCAGCCAGCAGCAGCAGGAGTGGACATCCCCCTT----- 229
Db      50 ProAlaLysArgProAlaAspAspThrAspLeuAsnProPheMetAspAspGluAla 69
QY      230 -----CGGCTCTCTGGTCCACCCAGTATGTGGTCCCATATTGCGAAGGAGGG 280
Db      70 ValAsnGluLysTyrProIleProGluSerAlaValGlyIleValIleGlyArgGlyGly 89
QY      281 GCCACCATCCGACATCAAAACAGACCCAGTCCAGATAGACGTGCATAGGAAGGAG 340
Db      90 SerGluIleGlnGlyIleGlnAlaLysAlaGlyCysArgValGlnMetSerProAspAla 109
QY      341 AACCGAGTGCAGCTGAAAAGCCATCAGTGTGCATCTCCACCCCTCGAGGCTCTCTCC 400
Db      110 AspProSerSerGlyValArgMetValThrLeuGluGlySerArgSerAsnValGluThr 129
QY      401 GCTTGTAAAGATCTTGGAGATTATGCATAAGAGGCTAAGACACCAAA----- 451
Db      130 AlaLysHisLeuIleAsnGluValAlaAlaArgSerGlnAsnProArgProGlnTyrGly 149
QY      452 ---ACGGCTGACAGGTTCCCTCAAGATCTCGGCCCATATAACTTTGAGGCGCTC 508
Db      150 PheProArgAlaGlnThrThrIleAspIleAlaIleProProAsnArgCysGlyLeuIle 169
QY      509 ATTGGCAAGGAGACCGAAGCTGAGAGAGGTAGAGCAAGATACCGAGACAAATACCC 568
Db      170 IleGlyLysSerGlyAspThrIleArgGlnLeuGlnLysSerGlyCysLysMetIle 189

```

```

QY      569 ATCTCCTCGTTGCAAGAC---CTTACCCCTTTACAACCCCTGAGAGGACCACCTACCTGTGAAG 625
Db      190 Leu-----ValGlnAspAsnGlnSerValSerAspGlnSerLysProLeuArgIleThr 207
QY      626 GGGGCCATCCAGAAATTTTGCAGGGCCGAGCAAGAAATAATGAAGAAAGTTTCGGGAGGCC 685
Db      208 GlyAspProGlnLysIleGluLeuAlaLysGlnLeuValAlaGluIleLeuAsnSerGly 227
QY      686 TATGAGATGATGTGGTGCATGCATGCTCTCACCTGATCCCTGCCTGAACCTGGCTGCT 745
Db      228 GlyAspGlyAsnGlySer-----GlyLeuGlnMetHis 240
QY      746 GTAGTCTTTTCCAGCTTTCATCCAGCGCAGTCCCGCGCTCCAGCAGCGTTTACTGGG 805
Db      241 AlaGly-----Gly 243
QY      806 GCTGCTCCTATAGTCTCTTTATGAGGCTCCCGAGCAGGAGATGGTGCAGGTGTTTATC 865
Db      244 GlyGlyGlyAlaSerAla-----ArgGlyGluValVal 255
QY      866 CCCGCCAGGAGTGGGGCCCATCATCGCAAGAGGGGAGCACATCAACACAGCTCTCC 925
Db      256 ProArgSerSerValGlyIleIleGlyLysGlnGlyAspThrIleLysArgLeuAla 275
QY      926 CGGTTTGCAGCGCTCCATCAAGATTGCACACCCCGAAACACCTGACTCCAAAGTTGCT 985
Db      276 MetGluThrGlyThrLysIleGlnPheLysProAspAspProSerThrProGluArg 295
QY      986 ATGTTATCATCACTGACCCCGCAGCCCAATCAAGGCTCAGGAGAGATATATGGC 1045
Db      296 CysAlaValIleMetGlyThrArgAspGlnIleThrArgAlaThrGluArgIleThrGlu 315
QY      1046 AAATCTCAAGGAGGAG-----AACTTCTTTTGGTCCCAAGGAG 1081
Db      316 LeuValLysLysSerThrMetGlnGlnGlyGlyGlyAsnValAlaGlyAlaMetVal 335
QY      1082 GAAGTGAAGCTGGAGACC---CACATAGTGTGCCAGCATCAGCAGCTGGCGGTCAATT 1138
Db      336 SerAsnGluAlaSerThrPheTyrMetSerValProAlaAlaLysCysGlyLeuValIle 355
QY      1139 GGCRAAGTGGAAAACCGTGAACGAGTTGCAGATTTGACGGCAGCTGAGGTGGTAGTA 1198
Db      356 GlyLysGlyGlyLuthrIleLysGlnIleAsnSerGluSerGlyAlaHisCysGluLeu 375
QY      1199 CCAGAGACCAACCCCTGTAGTGAACACGACGATCATCTGTG-----AAA 1243
Db      376 SerArgAspProThrGlyAsnAlaAspGluLysValPheValIleLysGlyLysArg 395
QY      1244 ATCATCGACATTTCTATCCAGTCAGATGCTCAACGAGAGATCCGAGAGATCTTGCC 1303
Db      396 AlaIleGluHis-----AlaLysHisLeuIleArgIleLysValGlyAspIleAlaPro 413
QY      1304 CAGGTTAAGCAGCATCAGAAAGGACAGAGTAAACGAGGCCCGCAGGAGGAAGTGA 1363
Db      414 AsnThrProPheArgAspAspSerAlaMetThrMetGlnThrGlnPhe----- 429
QY      1364 CCAGCCCTCTCTGTCCCTTNGAGTCCAGGACAAACACGGGCGAGAAATCGAGAGTGTCT 1423
Db      430 ---SerAlaProAlaGlnAsnAsnPheGlyGlyGlnGlnGlnTrpAsnProValAlaGln 448
QY      1424 CTCGCCGCGAGCCCTGAGATGAGTGGGAATCCGGG----- 1459
Db      449 IleProAlaAlaAlaGlnAsnProTyrGlnValGlyGlyTrpGlnGlnAsnSerValTyr 468
QY      1460 -----ACACNTGGCGCGCTGTAGATCAGTTTG----- 1489
Db      469 AlaGlnGlnThrAlaAlaProAlaAlaProTyrAlaAlaAlaGlyIleValGlnPro 488
QY      1490 -----CCCACTTGATTGAGAAAGATCTCCAGTGAGAAC 1525
Db      489 GlnGlnGlnValAlaTyrGlnGlnProGlnValValGlnGlnVal---GlnProAlaThr 507
QY      1526 CTGATCTNTCAGCCCAACACCCACC----- 1552

```

Db 508 ThrValAlaAlaThrThrThrProThrValAspProValThrGlyGluGlnAspTyrSer 527  
 QY 1553 ---CAATTGGCCCAACTGNTCCCTCGGGTG----- 1585  
 Db 528 AlaGlnTrpMetGluTyrTyrLysSerIleGlyAlaHisAspLysAlaGluAlaValGlu 547  
 QY 1586 -----TCAGAAATTNTAGCGCAAGGCACCTTTTAAACGTGGATTG 1624  
 Db 548 AlaGlnMetLysLysLysAlaGluAlaAlaAlaAlaAlaAlaValProGlyGlyLeu 567  
 QY 1625 TTAAAGAGCTCTCCAGGCCCAACCAAGAGGGTGGATCACCTCAG 1672  
 Db 568 -----IleGlnGlnMetProMetGlyMetAlaMetProGln 579  
 RESULT 12  
 138489  
 onconeural ventral antigen-1 - human  
 C:Species: Homo sapiens (man)  
 C:Date: 06-Sep-1996 #sequence\_revision 06-Sep-1996 #text\_change 05-Nov-1999  
 C:Accession: I38489  
 R:Buckanovich, R.J.; Posner, J.B.; Darnell, R.B.  
 Neuron 11, 657-672, 1993  
 A:Title: Nova, the Paraneoplastic Ri Antigen, is Homologous to an RNA-Binding Protein ar  
 A:Reference number: I38489; MUID:94000830; PMID:8398153  
 A:Accession: I38489  
 A:Status: Preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-510 <RES>  
 A:Cross-references: EMBL:U04840; NID:g440877; PIDN:AAA16022.1; PID:g440878  
 C:Genetics:  
 A:Gene: Nova-1

Alignment Scores:  
 Pred. No.: 4.92e-08 Length: 510  
 Score: 221.50 Matches: 111  
 Percent Similarity: 34.89% Conservative: 68  
 Best Local Similarity: 21.64% Mismatches: 173  
 Query Match: 7.12% Indels: 161  
 DB: 2 Gaps: 17

US-09-270-437D-5 (1-1708) x I38489 (1-510)

QY 125 GAGATGGCGCGCGAGGGGCTTT-----GGCTCTCGGGTCTAG 163  
 Db 9 GlnAsnGlyThrHisThrGlyValProIleAspLeuAspProProAspSerArgLysArg 28  
 QY 164 CCGCGCAGGGCTACCTGTGGCAGCGGGGCCCCAGCCAGCAGCAAGTGCAC--- 220  
 Db 29 ProLeuGluAlaProProGluAlaGlySerThrLysArgThrAsnThrGlyGluAspGly 48  
 QY 221 ---ATCCCTCTCGGCTCTGTGGTGGCCAGTATGTGGTCCATTATGGCAGGAG 277  
 Db 49 GlnTyrPheLeuLysValLeuIleProSerTyrAlaAlaGlySerIleGlyLysGly 68  
 QY 278 GGGGCCACCATCGGCACATCAAAACAGACCCAGTCCAGATAGCTGATAGGAAG 337  
 Db 69 GlyGlnThrIleValGlnGlnLysGlyGlyThrGlyAlaThrIleLysLeuSerLysLeu 88  
 QY 338 GAGAACGCA-----GGTGAGCTGAAAGCCATCAGTGTGCACTCCACC 382  
 Db 89 SerLysSerLysAspPheTyrProGlyThrThrGluArgValCysLeuIleGlnGlyThr 108  
 QY 383 CTGTAGGGCTGTCTCCCTCCT-----TGTAAGATGATCTTGGAGATTATGCAT 430  
 Db 109 ValGluAlaLeuAsnAlaValHisGlyPheIleAlaGluLysIleArgGluMetProGln 128  
 QY 431 AAAGAGGCTAAG----- 442  
 Db 129 AsnValAlaLysThrGluProValSerIleLeuGlnProGlnThrValAsnProAsp 148  
 QY 443 -----GACACCAAAACGGCTGACGAGGTTCCC 469  
 Db 476 ThrIleThrGlyThrProAlaAlaThrGlnAlaAlaGln 488

Db 149 ArgIleLysGlnThrLeuProSerSerProThrThrThrLysSerSerProSerAspPro 168  
 QY 470 CTG-----AAGATCTTGGCCCATATAACTTTGTAGGGCGT 505  
 Db 169 MetThrThrSerArgAlaAsnGlnValLysIleIleValProAsnSerThrAlaGlyLeu 188  
 QY 506 CTCATTGGCAAGGAGCAGCACTGAAAGAGGTAGAGCAAGATACCGACACAAAATC 565  
 Db 189 IleIleGlyLysGlyGlyAlaThrValLysAlaValMetGluGlnSerGlyAlaTrpVal 208  
 QY 566 ACCATCTCTCTGTTGCAAGACTTACCTTTACACCTGAGAGGACCATCAGTGTGAAG 625  
 Db 209 GlnLeuSerGlnLysProAsp---GlyIleAsnLeuGlnArgValValThrValSer 227  
 QY 626 GGGGCCATCGAGAATTGTTGCGAGGCGCAGGAGGAATAATGAGAAGTTCGGGAG--- 682  
 Db 228 GlyGluProGluGlnAsnArgLysAlaValGluLeuIleGlnLysIleGlnGluAsp 247  
 QY 683 -----GCCATATGAGAAAT----- 694  
 Db 248 ProGlnSerGlySerCysLeuAsnIleSerTyrAlaAsnValThrGlyProValAlaAsn 267  
 QY 695 -----GATGTGGTGGTCCCATG 709  
 Db 268 SerAsnProThrGlySerProTyrAlaAsnThrAlaGluValLeuProThrAlaAla 287  
 QY 710 AGCTCTCACCTGATCCCTGCGCTGACCTGCTGCTAGGTCTTTTCCCA----- 760  
 Db 288 AlaAlaGlyLeuLeuGlyHisAlaAsnLeuAlaGlyValAlaAlaPheProAlaValLeu 307  
 QY 760 ----- 760  
 Db 308 SerGlyPheThrGlyAsnAspLeuValAlaIleThrSerAlaLeuAsnThrLeuAlaSer 327  
 QY 761 -----GCTTCATCCAGCGCA 775  
 Db 328 TyrGlyTyrAsnLeuAsnThrLeuGlyLeuGlyLeuSerGlnAlaAlaIleThrGlyAla 347  
 QY 776 GTCCCGCGGCTCCCGCAGCAGGTACTGGGGTGTCTCCCTATAGTCTCTTATGCAAGCT 835  
 Db 348 LeuAlaAlaAlaAlaAlaSerAlaAsnProAlaAlaAlaAlaAlaAsnLeuAlaThr 367  
 QY 836 CCGCAGCAGCAGAGTGGTGCAGGTGTTATCCCGCCCGCAGCAGTGGCGCCATCGGC 895  
 Db 368 TyrAlaSerGlu-----AlaSerAlaSerGlySerThrAlaGly 380  
 QY 896 AAGAGGGGCGCAGCACATCAAAACAGCTCTCCCGTTTCCAGCGCTCCATCAAGATTGCA 955  
 Db 381 GlyThrAla-----GlyThrPheAlaLeuGlySerLeuAlaAlaAla 394  
 QY 956 CCACCGAAACACCTGACTCCAAAGTTCGTATGTTATCATCTAGCTGCGCCAGAGGCC 1015  
 Db 395 -----ThrAlaAlaThrAsnGly 400  
 QY 1016 CAATTCAAGCTCAGGAGAAATCTATGCAAACTC-----AAGAGGAGAACTTC 1066  
 Db 401 TyrPheGlyAlaAlaSerProLeuAlaAlaSerAlaIleLeuGlyThrGluLysSerThr 420  
 QY 1067 TTTGGTCCCAGGAGAAAGTGAAGCTGGAGACCCACATAGCTGTGCGCAGCATCAGAGCT 1126  
 Db 421 AspGlySerLysAspValValGlu-----IleAlaValProGluAsnLeuVal 436  
 QY 1127 GGGCGGCTATTCCGCAAGGTGAAAAACGGTGAACAGTTCGAGAATTTCAGCGCAGCT 1186  
 Db 437 GlyAlaIleLeuGlyLysGlyLysThrLeuValGluTyrGlnGluLeuThrGlyAla 456  
 QY 1187 GAGTGTGTAGTACCAAGA-----GACCAGACCCCTGATGAGAACGACCGAGTTCATCGTG 1240  
 Db 457 ArgIleGlnIleSerLysLysGlyGluPheValProGlyThrArgAsnArgLys---Val 475  
 QY 1241 AAAATCATCGACATTTCTATCCAGTCAGATGGCTCAA 1279  
 Db 476 ThrIleThrGlyThrProAlaAlaThrGlnAlaAlaGln 488





```

QY 536 AAGGTAGACAGATACCGAGACAAAATCACCATCTCTCTGTTGCAAGACCTTACCCCTT 595
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
145 SerPheIleGluSerLysAlaGlyIleLysIleSerProLeu---AspAsnThrPhe 163
QY 596 TACAC----CCTGAGAGGACCATCACTGTGAAGGGGCCCATCGAGAAATTGTTGCAGGGCC 652
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
164 TyrGlyLeuSerAspArgLeuValThrLeuSerGlyThrPheGluGluGlnMetArgAla 183
QY 653 GAGCAGGAAATATGAGAAAGTTCCGGAGGCGCTATGAGATGATGTGGCTGCCATGAGC 712
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
184 IleAspLeuIleLeuAlaLysLeuThr-----GluAspAspHisTyrSerGlnAsn 200
QY 713 TCTCACTGATCCCTGGCTGAACCTGGCTGCTGTAGTCTTTCCACGCTTCATCCAGC 772
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
201 ValHis----- 202
QY 773 GCAGTCCCGCGCTCCCGAGCAGGTACTGGGGCTGCTCCCTATAGCTTCCTTTATGCAG 832
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
203 -----SerProTyrSerTyrAlaAlaGlyTyrAsnSerValAsnTyr 216
QY 833 GCTCCCGAG-----CAGGAGATGGTG 853
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
217 AlaProAsnGlySerGlyLysTyrGlnAsnHisLysGluGluAlaSerThrThrVal 236
QY 854 CAGGTGTTTATCCCGCCAGGCGAGTGGCGCCATCATCGCAAGAGGGCGAGCACATC 913
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
237 ThrIleGlyValAlaAspGluHisIleGlyLeuValLeuGlyArgGlyGlyArgAsnIle 256
QY 914 AAACAGCTCTCCCGGTTGCCAGCGCTCCATCAAGATTGCACCA----- 958
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
257 MetGluIleThrGlnMetThrGlyAlaArgIleLysIleSerAspArgGlyAspPheMet 276
QY 959 CCCGAACACCTGACTCCCAAGTTCGTATGGTTATCATCACTGGACCGCCAGAGGCCCAA 1018
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
277 SerGlyThrThrAspArgLysValSer-----IleThrGlyProGlnArgAlaIle 293
QY 1019 TTCAAGGCTCAGGGAAGAATCTATGGCAAACTC 1051
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
294 GlnGlnAlaGluThrMetIleLysGlnLysVal 304

```

Search completed: July 13, 2004, 12:13:47  
Job time : 53 secs



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: July 13, 2004, 11:52:36 ; Search time 19.5 seconds

(without alignments)  
9121.623 Million cell updates/sec

Title: US-09-270-437D-5  
Perfect score: 3110  
Sequence: 1 agggacgcgcgcacgcgc.....atttccttcaggttttaaaa 1708

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 283362

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp  
-Q=/cgn2\_1/USPRO.spool\_P/US09270437/runat\_13072004\_121921\_9590/app\_query.fasta\_1.1863  
-DB=SwissProt 42 -QFMT=fastan -SUFFIX=resp -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pco -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09270437@cgn\_1\_1\_24@runat\_13072004\_121921\_9590 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID           | Description         |
|------------|-------|-------------|--------|--------------|---------------------|
| 1          | 264   | 8.5         | 572    | 1 FUB3_HUMAN | Q96124 homo sapien  |
| 2          | 248   | 8.0         | 651    | 1 FUB1_MOUSE | Q91WJ8 mus musculus |
| 3          | 247.5 | 8.0         | 643    | 1 FUB1_HUMAN | Q96ae4 homo sapien  |
| 4          | 236.5 | 7.6         | 492    | 1 NOA2_HUMAN | Q9unw9 homo sapien  |
| 5          | 235.5 | 7.6         | 769    | 1 2BP2_CHICK | Q8uvd9 gallus galli |
| 6          | 234.5 | 7.5         | 707    | 1 FUB2_HUMAN | Q92945 homo sapien  |
| 7          | 231   | 7.4         | 474    | 1 NOA1_RAT   | Q80wa4 rattus norv  |
| 8          | 228.5 | 7.3         | 721    | 1 FUB2_RAT   | Q99pf5 rattus norv  |
| 9          | 221.5 | 7.1         | 356    | 1 PCB1_HUMAN | Q15365 homo sapien  |
| 10         | 221.5 | 7.1         | 493    | 1 NOA1_MOUSE | Q9jkn6 mus musculus |
| 11         | 221.5 | 7.1         | 510    | 1 NOA1_HUMAN | P51513 homo sapien  |
| 12         | 218.5 | 7.0         | 356    | 1 PCB1_RABIT | O19048 oryctolagus  |
| 13         | 217.5 | 7.0         | 403    | 1 PCB4_HUMAN | P57723 homo sapien  |
| 14         | 214.5 | 6.9         | 403    | 1 PCB4_MOUSE | P57724 mus musculus |
| 15         | 209   | 6.7         | 606    | 1 Y475_ARATH | P58223 arabidopsis  |
| 16         | 209   | 6.7         | 1268   | 1 VGLN_ARATH | Q00341 homo sapien  |
| 17         | 208.5 | 6.7         | 339    | 1 PCB3_HUMAN | P57721 homo sapien  |
| 18         | 207.5 | 6.7         | 339    | 1 PCB3_MOUSE | P57722 mus musculus |

|    |       |     |      |               |                     |
|----|-------|-----|------|---------------|---------------------|
| 19 | 203.5 | 6.5 | 1270 | 1 VGLN_CHICK  | P81021 gallus galli |
| 20 | 199   | 6.4 | 365  | 1 PCB2_HUMAN  | Q15366 homo sapien  |
| 21 | 197.5 | 6.4 | 463  | 1 ROK_HUMAN   | Q07244 homo sapien  |
| 22 | 197.5 | 6.4 | 463  | 1 ROK_RABIT   | O19049 oryctolagus  |
| 23 | 194   | 6.2 | 362  | 1 PCB2_MOUSE  | Q61990 mus musculus |
| 24 | 191   | 6.1 | 413  | 1 PBP2_YEAST  | P38151 saccharomyc  |
| 25 | 182.5 | 5.9 | 1222 | 1 S160_YEAST  | P06105 saccharomyc  |
| 26 | 178   | 5.7 | 381  | 1 YBD2_YEAST  | P38193 saccharomyc  |
| 27 | 170.5 | 5.5 | 606  | 1 TDRH_HUMAN  | Q9yzw6 homo sapien  |
| 28 | 164   | 5.2 | 784  | 1 YAV2_XANCV  | P14728 xanthomonas  |
| 29 | 153   | 4.9 | 2805 | 1 MAPA_HUMAN  | P78559 homo sapien  |
| 30 | 144.5 | 4.6 | 1336 | 1 W146_HUMAN  | Q9C0J8 homo sapien  |
| 31 | 143   | 4.6 | 648  | 1 GLUT0_WHEAT | P10387 triticum ae  |
| 32 | 142   | 4.6 | 2142 | 1 BAT2_HUMAN  | P48634 homo sapien  |
| 33 | 141.5 | 4.5 | 839  | 1 GLUT5_WHEAT | P10388 triticum ae  |
| 34 | 139.5 | 4.5 | 1152 | 1 MAP4_HUMAN  | P27816 homo sapien  |
| 35 | 139.5 | 4.5 | 1181 | 1 HAIR_RAT    | P97609 rattus norv  |
| 36 | 138.5 | 4.5 | 992  | 1 EBN6_EBV    | P03204 Epstein-Bar  |
| 37 | 135   | 4.3 | 639  | 1 ZIC5_HUMAN  | Q96t25 homo sapien  |
| 38 | 133   | 4.3 | 838  | 1 GLUT4_WHEAT | P08489 triticum ae  |
| 39 | 132.5 | 4.3 | 1820 | 1 STR9_HUMAN  | Q9a2p6 homo sapien  |
| 40 | 132.5 | 4.3 | 2716 | 1 OSA_DROME   | Q8in94 drosophila   |
| 41 | 131   | 4.2 | 931  | 1 MPTA_HUMAN  | Q969v6 homo sapien  |
| 42 | 131   | 4.2 | 1182 | 1 HAIR_MOUSE  | Q61645 mus musculus |
| 43 | 131   | 4.2 | 5147 | 1 PCLO_HUMAN  | Q9y6v0 homo sapien  |
| 44 | 130.5 | 4.2 | 1461 | 1 IE18_PRTIF  | P11675 pseudorabie  |
| 45 | 128.5 | 4.1 | 1729 | 1 TABP_HUMAN  | Q9C0C2 homo sapien  |

#### ALIGNMENTS

RESULT 1  
FUB3\_HUMAN  
ID FUB3\_HUMAN STANDARD; PRT; 572 AA.  
AC Q96124; Q92946; Q9BVB6;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Far upstream element binding protein 3 (FUSE binding protein 3).  
GN FUBP3  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RX MEDLINE=97094955; PubMed=8940189;  
RA Davis-Smyth T., Duncan R.C., Zheng T., Michelotti G., Levens D.;  
RT "The far upstream element-binding proteins comprise an ancient family  
of single-strand DNA-binding transactivators.";  
RL J. Biol. Chem. 271:31679-31687(1996).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
RC TISSUE=Cervix and Placenta;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,  
RA Klausner R.D., Collins S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ustun T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalios D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length





```

QY 1160 AACGAGTTCAGAAATTGACGGCAGCTGAGGTGTAGTACCAGAGCAGACCCCTGAT 1219
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 378 LysSerIleasnGlnGlnSerGlyAlaHisValGluLeuGlnArgAsnProProAsn 397
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1220 GAGAACGACGACCGATCATCGTGAATAATCATCGGACATTTCTATGCCAGTCAGTGCACAA 1279
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 398 SerAsp-----ProAsnLeuArgArg 404
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1280 CGGAAGATCCGACATCTGCGCCAGGTTAAGCAGCAGCATCAGAG----- 1327
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 405 PheThrIleArgGlyValProGlnGlnIleGluValAlaArgGlnLeuIleaspGluLys 424
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1328 -----GGACAGAGTAACAGCCCGCCAGCGCA 1351
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 425 ValGlyGlyThrAsnLeuGlyAlaProGlyAlaPheGlyGlnSerProPheSerGln--- 443
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1352 CGAGGAGTGCACCGCCCTCCTGCTCCTTGAGTCCAGACACACAGCGGCAGAAAT 1411
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 444 -----ProProAlaProProHisGlnAsnThrPheProProArgSer 457
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1412 CGAGAGTGTGCTCTCCCGCGCAGCCTCGAGAAATGAGTGGGAATCCG----- 1457
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 458 SerGlyCys-PheProAsnMetAlaAlaLysValAsnGlyAsnProHisSerThrProVa 477
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1458 -----GGACACNTGGCGCGCGCTGTAGATCA 1483
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 477 lSerGlyProProAlaPheLeuThrGlnGlyTrpGlySerThrTyTGlnAlaTrpGlnGl 497
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1484 GGTTCGCCACTTGATGAGAAAGATGTTCCAGTGAGGAACTCGATCNTCAGCCCA 1543
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 497 n-----ProThrGlnValProSerGlnGlnSerGlnProGlnSerSerGlnProAs 515
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1544 ACACCCACCAATGG 1559
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 515 nTyrSerLysAlaTrp 520
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 2
FUBI_MOUSE STANDARD; PRT; 651 AA.
ID FUBI_MOUSE AC Q91WJ8; Q8C0Y8;
RX MEDLINE=22388257; PubMed=12477932;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Far upstream element binding protein 1 (FUSE binding protein 1) (FBP).
GN FUBP1 OR D3ERTD330E.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Retina;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
```

```

RN [2]
RP SEQUENCE OF 1-286 FROM N.A. (ISOFORM 2).
RC STRAIN=CS7BL/6J; TISSUE=Head;
RX MEDLINE=22354683; PubMed=12466851;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojohori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schirali L.M., Kanapin D., Mateuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grømond S., Guetinch S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kutochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Mutsaers L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Nunata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramchandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wallestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang L., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carinci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Komoto H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
CC -!- FUNCTION: Regulates MYC expression by binding to a single-stranded
CC far-upstream element (FUSE) upstream of the MYC promoter. May act
CC both as activator and repressor of transcription (By similarity).
CC -!- SUBUNIT: Interacts with SIAH3p1/IR and JTV1 (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q91WJ8-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q91WJ8-2; Sequence=VSP_008322;
CC Note=No experimental confirmation available;
CC -!- PTM: Ubiquitinated. This targets the protein for proteasome-
CC mediated degradation (By similarity).
CC -!- SIMILARITY: Contains 4 KH domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; BC014763; AAH14763.1; -;
CC EMBL; AK029458; SACC6457.1; -;
CC MGD; MGI:1196294; D3ErtD330e.
CC MGD; MGI:1924642; 9530027K12Rik.
CC InterPro; IPR008160; Collagen.
CC InterPro; IPR004087; KH dom.
CC InterPro; IPR004088; KH_type_1.
CC Pfam; PF00013; KH; 4.
CC SMART; SM00322; KH; 4.
CC PROSITE; PS50084; KH TYPE_1; 4.
CC Transcription regulation; Trans-acting factor; Nuclear protein;
CC DNA-binding; Repeat; Ubl conjugation; Alternative splicing.
CC DOMAIN 96 160 KH 1.
CC DOMAIN 181 247 KH 2.
CC DOMAIN 271 335 KH 3.
CC DOMAIN 372 439 KH 4.
CC
```



QY 1663 CACACC-----TCAGTGGGAAGAAATAAAATTCCTTCA 1698  
 Db 578 TyrTrpLysAlaTrpGluGluTyrTyrLysLysMetGlyGlnAlaValProAlaProAla 597  
 QY 1699 GGT 1701  
 Db 598 Gly 598  
 RESULT 3  
 ID FUBI\_HUMAN STANDARD; PRT; 643 AA.  
 AC Q96AE4; Q12828;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Far upstream element binding protein 1 (FUSE binding protein 1) (FBP)  
 DE (DNA helicase V) (HDH V).  
 GN FUBP1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1), SEQUENCE OF 191-193; 203-206;  
 RP 272-279; 284-290; 300-314; 321-328; 394-397; 409-411; 430-438 AND  
 RP 440-443, AND FUNCTION.  
 RC TISSUE=Leukemia;  
 RX MEDLINE=9417091; PubMed=8125259;  
 RA Duncan R., Bazar L., Micheletti G., Tomonaga T., Kruttsch H.,  
 RA Avigan M., Levens D.;  
 RT "A sequence-specific, single-strand binding protein activates the far  
 RT upstream element of c-myc and defines a new DNA-binding motif.";  
 RL Genes Dev. 8:465-480(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RC TISSUE=Brain;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Krausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
 RA Altshul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gumaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [3]  
 RP SEQUENCE OF 45-64; 133-145; 271-283; 308-321; 379-386; 414-424 AND  
 RP 430-439.  
 RX MEDLINE=21127960; PubMed=11222755;  
 RA Vindigni A., Ocham A., Triolo G., Falaschi A.;  
 RT "Identification of human DNA helicase V with the far upstream  
 RT element-binding protein.";  
 RL Nucleic Acids Res. 29:1061-1067(2001).  
 RN [4]  
 RP PARTIAL SEQUENCE, AND MASS SPECTROSCOPY.  
 RX MEDLINE=22166132; PubMed=12176931;  
 RA Rappesilber J., Ryder U., Lamond A.I., Mann M.;  
 RT "Large-scale proteomic analysis of the human spliceosome.";  
 RL Genome Res. 12:1231-1245(2002).  
 RN [5]

RP INTERACTION WITH SIAHBP1.  
 RX MEDLINE=20337922; PubMed=10882074;  
 RA Liu J., He L., Collins I., Ge H., Libutti D., Li J., Egly J.-M.,  
 RA Levens D.;  
 RT "The FBP interacting repressor targets TFIID to inhibit activated  
 RT transcription.";  
 RL Mol. Cell 5:331-341(2000).  
 RN [6]  
 RP INTERACTION WITH JTV1, UBIQUITINATION, AND PROTEASOME-MEDIATED  
 RP DEGRADATION.  
 RX MEDLINE=22716800; PubMed=12819782;  
 RA Kim M.J., Park B.-J., Kang Y.-S., Kim H.J., Park J.-H., Kang J.W.,  
 RA Lee S.W., Han J.M., Lee H.-W., Kim S.;  
 RT "Downregulation of FUSE-binding protein and c-myc by tRNA synthetase  
 RT cofactor p38 is required for lung cell differentiation.";  
 RL Nat. Genet. 34:330-336(2003).  
 RN [7]  
 RP STRUCTURE BY NMR OF 278-447 IN COMPLEX WITH SINGLE STRANDED FUSE DNA.  
 RX MEDLINE=21864720; PubMed=11875576;  
 RA Braddock D.T., Louis J.M., Baber J.L., Levens D., Clore G.M.;  
 RT "Structure and dynamics of KH domains from FBP bound to  
 RT single-stranded DNA.";  
 RL Nature 415:1051-1056(2002).  
 CC -!- FUNCTION: Regulates MYC expression by binding to a single-stranded  
 CC far-upstream element (FUSE) upstream of the MYC promoter. May act  
 CC both as activator and repressor of transcription.  
 CC -!- SUBUNIT: Interacts with SIAHBP1/FIR and JTV1.  
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing; Named isoforms=2;  
 CC Name=1;  
 CC IsoId=Q96AE4-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=Q96AE4-2; Sequence=VSP\_008321;  
 CC Note=NO experimental confirmation available;  
 CC -!- PTM: Ubiquitinated. This targets the protein for proteasome-  
 CC mediated degradation.  
 CC -!- SIMILARITY: Contains 4 KH domains.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U05040; AAA17976.2; -.  
 CC EMBL; BC017247; AAA17247.1; -.  
 CC PDB; 1J4W; 06-MAR-02.  
 CC Genew; HGNC:4004; FUBP1.  
 CC MIM; 603444; -.  
 CC InterPro; IPR008160; Collagen.  
 CC InterPro; IPR004087; KH dom.  
 CC InterPro; IPR004088; KH\_type\_1.  
 CC Pfam; PF00013; KH; 4.  
 CC SMART; SM00322; KH; 4.  
 CC PROSITE; PS50084; KH TYPE 1; 4.  
 CC Transcription regulation; Trans-acting factor; Nuclear protein;  
 CC DNA-binding; Repeat; Ub1 conjugation; Alternative splicing;  
 CC 3D-structure.  
 FT DOMAIN 99 163 KH 1.  
 FT DOMAIN 184 250 KH 2.  
 FT DOMAIN 274 338 KH 3.  
 FT DOMAIN 375 442 KH 4.  
 FT DOMAIN 13 26 GLY-RICH.  
 FT DOMAIN 348 395 GLY-RICH.  
 FT DOMAIN 449 559 PRO-RICH.  
 FT VARSPLIC 642 643 GQ -> CRFDPASIELAL (in isoform 2).  
 FT /FTId=VSP\_008321.  
 FT CONFLICT 96 96 Q -> QS (IN REF. 1).  
 FT SEQUENCE 643 AA; 67473 MW; 086D4EAA0ACF807B CRC64;

## Alignment Scores:

Pred. No.: 14e-08 Length: 643  
 Score: 247.50 Matches: 140  
 Percent Similarity: 37.80% Conservative: 97  
 Best Local Similarity: 22.33% Mismatches: 202  
 Query Match: 7.96% Indels: 189  
 DB: 1 Gaps: 32

US-09-270-437D-5 (1-1708) x FUBI\_HUMAN (1-643)

```

Qy 98 CCCGAT---GAGCAGATAGACAGGAGCTGAGATGGCGCGGCGGCTTTGGCTCT 154
Db 73 ProAspAlaLysValAla-----ProGlnAsn-----AspSerPheGlyThr 87
Qy 155 CGGGGTGAGCCCGCCAGGCTCACCTGTGGCAGCGGGGCGCCAGCCAGCAGCAAA 214
Db 88 GlnLeuProMetHis-----GlnGlnGlnArg 97
Qy 215 GTGGACATCCCGCTTCGGCTCCGTGGTCCCGCCAGCAGTATGTGGTGCCATTATTGGCAAG 274
Db 98 SerValMetThrGluGluTyrLysValProAspGlyMetValGlyPheIleIleGlyArg 117
Qy 275 GAGGGGCCACCATCCGCAACATCACAACAGACCCAGCTCCAGATAGACGTGCATAG 334
Db 118 GlyGlyGluGlnIleSerArgIleGlnGlnGluSerGlyCysLysIleGlnIle---Ala 136
Qy 335 AAGGAGAACCCAGGTGCAGCTGAAAAAGCCATCAGTGTGCACCTCCACCCCTCAGGGCTGC 394
Db 137 ProAspSerGlyGlyLeuProGluArgSerCysMetLeuThrGlyThrProGluSerVal 156
Qy 395 TCCTCCGCTTGTAAAGATGCTTGGAGATATGCAATAA-----GAGCTAAGACACC 448
Db 157 GlnSerAlaLysArgLeuLeuAspGlnIleValGluLysGlyArgProAlaProGlyPhe 176
Qy 449 AAAAGCGCTCAC-----GAGGTCCCTGAGCATCTCGCCCATTAATCTTGTA 499
Db 177 HisHisGlyAspGlyProGlyAsnAlaValGlnGluIleMetIleProAlaSerLysAla 196
Qy 500 GGCGGTCTCATTCGCAAGAGCAGCGAACCCTGAGAGGTAGAGCAAGATACCCAGACA 559
Db 197 GlyLeuValIleGlyLysGlyGlyGluThrIleLysGlnLeuGlnArgAlaGlyVal 216
Qy 560 AAATACCATCTCCCTGCTGTCAGACCTTACCTTTACAACTCAGAGGACC----- 613
Db 217 LysMetValMet-----IleGlnAsp-----GlyProGlnAsnThrGlyAla 230
Qy 614 -----ATCACTGTGAAGGGGGCCATCGAGAATTGTGAGGGCCCGAGCAAAATA 664
Db 231 AspLysProLeuArgIleThrGlyAspProTyrLysValGlnGlnAlaLysGluMetVal 250
Qy 665 ATGAAGAAGTTCGGAG-----GCCTATGAGATGATGTGCTGCCATG 709
Db 251 LeuGluLeuIleArgAspGlnGlyPheArgGluValArgAsnGluTyrGly----- 268
Qy 710 AGCTCTCACCTGATCCCTGCCTGACCTGAGCTGCTAGGTCTTTCCACGCTTCATCC 769
Db 268 ----- 268
Qy 770 AGGCGAGTCCCGCGCTCCAGCAGCGTTATGTGGGTGCTCCCTATAGCTCCTTTATG 829
Db 269 -----SerArgIleGlyGly----- 273
Qy 830 CAGGCTCCCGAGCAGGATGTGTGAGGTGTTATCCCGCCCGCCAGGCGAGTGGCGCCATC 889
Db 274 -----AsnGluGlyIleAspValProIleProArgPheAlaValGlyIleVal 289
Qy 890 ATGGCAAGAGCGGCGAGCAGCATCAAAACAGCTCTCCCGGTTTCCAGCGGCTCCCATCAAG 949
Db 290 IleGlyArgAsnGlyGluMetIleLysLysIleGlnAsnAspAlaGlyValArgIleGln 309
Qy 950 ATTGCACACCCGAA-----ACACTGACTCCAAAGTTCGTATGTTATCATCATCTCGA 1003
Db 310 PheLysProAspAspGlyThrThrProGlu-----ArgIleAlaGlnIleThrGly 326

```

```

Qy 1004 CGCCGAGAG---GCCCAATTCAAGGCTCAG----- 1030
Db 327 ProProAspArgCysGlnHisAlaAlaGluIleIleThrAspLeuLeuArgSerValGln 346
Qy 1031 -----GGAGAATCTATGCAACTCAAGGAGGAG 1060
Db 347 AlaGlyAsnProGlyGlyProGlyProGlyArgGlyArgGlyGlyGlyGlyGlyGly 366
Qy 1061 AACTTCTTTGTCCTCCCAAGGAGGAGTGAAGCTGGAGACCCACATACCTGTGCCAGCATCA 1120
Db 367 TrpAsnMetGlyProProGlyGlyLeuGln---GluPheAsnPheIleValProThrGly 385
Qy 1121 GCAGCTGCGCGGTCATTGGCAAGGTGGAAAACCGTGAAACGATGTCAGATTTCGACG 1180
Db 386 LysThrGlyLeuIleIleGlyLysGlyGlyGlyThrIleLysSerIleSerGlnGlnSer 405
Qy 1181 GCAGCTGAGGTGTGTAGTACCAAGAGACGACGACCCCTGATGAGAACACCAAGGTC---ATC 1237
Db 406 GlyAlaArgIleGluLeuGlnArgAsnProProAsnAlaAspProAsnMetLysLeu 425
Qy 1238 GTGAATATCATCGACAT-----TTCTATCCAGTCAGATGGCTCAACGGAAG 1285
Db 426 PheThrIleArgGlyThrProGlnGlnIleAspTyrAlaArgGlnLeuIleGluLys 445
Qy 1286 AT-----CGAGACATCCT---GGCCAGGTAAAGCAGCAGCATCAGAAGGAGCAGAG 1335
Db 446 IleGlyGlyProValAsnProLeuGlyProProValProHisGlyProHisGlyVal--- 464
Qy 1336 TAACAGCGCCAGGACCGA-----GGAAGTACGACGACCCCTC----- 1373
Db 465 ---ProGlyPro---HisGlyProProGlyProGlyProGlyProGlyProMetGlyPro 483
Qy 1374 -----CTGTCTCCCTNGAGTCCAGGA----- 1394
Db 483 TyrAsnProAlaProTyrAsnProGlyProProGlyProAlaProHisGlyProProAlaP 503
Qy 1395 -----CAACAACGGCGAGAAATCG 1413
Db 503 roTyrAlaProGlnGlyTyrGlyAsnAlaTyrProHisTyrGlnGlnAlaProProA 523
Qy 1414 AGAGTGTCTCTCCCGCGAGGCTGAGATGAGTGGGATCCGGACACNTGGCGGG 1473
Db 523 sp-----ProAlaLysAlaGlyThrAspProAsnSerAla-AlaTyrAlaAla 538
Qy 1474 CTGTAGATCAGGTTTGCCACTTGTGATGAGAAAGATGTTCCAGTGAGGAACCTGATCTN 1533
Db 539 Tyr-----TyrAlaHisTyrTyrGlnGlnAla----- 548
Qy 1534 TCAGCCCAACACCCACCCCAATGGC---CCAACTGNTGTCCTCGGGGTGTCAGA 1590
Db 549 GlnProProAlaAlaProAlaGlyAlaProThrThrGlnThrAsnGlyGlnGly 568
Qy 1591 AATTNTAGCCAAAGGCACTTTTAAACGTGGATGTTTAAAGAGCTCTCCAGGCCCCACC 1650
Db 569 Asp-----GlnGlnAsnProAlaProAla 576
Qy 1651 AAGAGGTGATCAGACAC-----TCAGTGGGAGAAAAATA 1686
Db 577 GlyGlnValAspTyrThrLysAlaTyrGluGluTyrTyrLysLysMetGlyGlnAlaVal 596
Qy 1687 AAATTTCTTCAGGT 1701
Db 597 ProAlaProThrGly 601

```

## RESULT 4

NOA2 HUMAN STANDARD; PRT; 492 AA.  
 ID AC Q9UNW9: O43267; Q9UEAL;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE RNA-binding protein Nova-2 (Neuro-oncological ventral antigen 2)





Db 227 yGluGluGlnIleAsnLysIleGlnGlnAspSerGlyCysLeuValGlnIle---SerPr 246  
Qy 337 GGAGAGCGAGGTGCGAGCTGAAAGAACCATCACTAGTGTGCACTCCACCCCTGAGGGCTGCTC 396  
Db 246 oAspSerGlyGlyLeuProGluArgSerValSerLeuThrGlySerProGluAlaValG1 266  
Qy 397 CTCGGCTTGAAGATGATCTTGAGATTATG----- 427  
Db 266 nLysAlaLeuLeuMetLeuAspIleValSerArgGlyArgGlyGlyProProGlyG1 286  
Qy 428 ----CATAAAGAGCTTAAG---GACACCAAAACGGCTGACAGGTTCCCTCAAGATCCT 480  
Db 286 nPheHisAspTyrAlaAsnGlyGlnAsnGlyThrValGlnGlu-----IleMe 302  
Qy 481 GGCCCATAAATCTTGTAGGGGCTCTCATTTGCAAGGAGGAGCGAACTGAAGAGGT 540  
Db 302 tileProAlaGlyLysAlaGlyLeuValIleGlyLysGlyGlyLeuThrIleLysGlnIle 322  
Qy 541 AGAGCAAGATACCGAGACAAATAATCACCATC-----TCTCGTTGCAAGACCTTACCCT 594  
Db 322 uGlnGluArgAlaGlyValLysMetIlePheIleGlnAspGlySerGlnAsnThrAsnVa 342  
Qy 595 TTACACACCTCGAGAGGACCATCACTGTGAAGGGGGCCATCGAGAATTGTTCAGGGCCGA 654  
Db 342 lAspLysProLeuArgIleIleGlyAspProTyrLysValGlnGlnAlaCysGluMetVa 362  
Qy 655 GCAGGAATATGAAGAAAGTTCGGAGGCTATGAGAAATGATGCTGCTGCCATGAGCTC 714  
Db 362 lMetAspIleLeuArgGluArgAspGlnGlyGlyPheGlyAspArgAsnGluTyrGlySe 382  
Qy 715 TCACCTGATCCCTGGCTGAGACCTGCTGCTGTAGGTCTTTTCCCGAGCTTCACCGAGC 774  
Db 382 rArgIleGlyGlyGlyIleAsp----- 389  
Qy 775 AGTCCCGCCGCTCCCGAGCGCTTACTGGGGCTGCTCCCTATAGTCTCTTTATGAGGC 834  
Db 390 -ValPro----- 391  
Qy 835 TCCGAGCAGGAGATGCTGAGGTGTTATTCCTCCGCCCGAGCGAGTGGGCCCATCATCGG 894  
Db 392 -----ValProArgHisSerValGlyValValIleG1 402  
Qy 895 CAAGAGGGGCGACACATCAACAGCTCTCCCGTTTGCAGCGCTCCATC----- 946  
Db 402 yArgSerGlyGluMetIleLysIleGlnAsnAspAlaGlyValArgIleGlnPheIy 422  
Qy 947 -----AAGATTGCA-----CCACCCGAAACACC 969  
Db 422 sGlnAspAspGlyThrGlyProGluLysIleAlaHisIleMetGlyProProGluArgCy 442  
Qy 970 TGACTCCAAAGTTCGTATGTT-----ATCATCACTGGACCGCCAGA 1011  
Db 442 sGluHisAlaAlaArgIleIleAsnAspLeuLeuGlnSerLeuArgSerGlyProPro-- 461  
Qy 1012 GGCCCAATT-----CAAGGCTCAGGGAAGATCTA 1041  
Db 462 GlyProProGlyHisGlyMetProProGlyArgGlyArgGlyArgGlyGlnGlyIle 481  
Qy 1042 TGGCAAACTCAAGGAGGAGAACTTCTTTGTTGCCAAGAGGAAAGTGAAGCTGGAGCCCA 1101  
Db 482 TrpGlyProProGlyGlyGluMetThrPhe----- 491  
Qy 1102 CATACGTGTGCAGCATCAGACGTGCGGGGTCATTGGCAAGAGTGGAAAAACGGTGAA 1161  
Db 492 ---Ser-IleProThrHisLysCysGlyLeuValIleGlyArgGlyGlyGluAsnVally 510  
Qy 1162 CGAGTTGCAGAAATTTGACGGCAGCTGAGGTGTAGTACCAAGAGACACAGACCCCT----- 1216  
Db 510 sAlaIleAsnGlnGlnArgGlyAlaPheValGluIleSerArgGlnLeuProProAsnG1 530  
Qy 1217 -GATGAGAACGACAGGTCATCGTGAATAATCATCGGACATTTCTATCCAGTCAGATGGC 1275  
Db 530 yAspProAsnPhelysLeuPheIle----- 538

Qy 1276 TCAACGGGAAGATCCGAGACATCTCTGGCCAGGTTAAGCAGCAGCATCAGAGGACAGAG 1335  
Db 539 -----IleArgGlySerProGlnGlnIleGluHisAlaLysGln----- 551  
Qy 1336 TAACAGGCCCGCAGGACGAGAGTGAACAGCCCTCCCTGTCCCTTNGAGTCCAGGAC 1395  
Db 552 -----ProIleGluGluLysIleGluGlyProLeuCysPro----- 563  
Qy 1396 AACACGGGCGAGAATCGAGAGTGTCTCTCCCGGCGAGGCTGAGAATGAGTGGGAATC 1455  
Db 563 ----- 563  
Qy 1456 CGGACACNTGGCGGGGCTGTAGATCAGGTTCGCCACTTTCATTGAGAAAGATGTTCCA 1515  
Db 564 -----ValGlyProGlyProGlyProGlyProPro----- 574  
Qy 1516 GTGAGGAACCTGATCTNTCAGCCCAACACCCACCACTGCCCACACTGNTGCC 1575  
Db 575 -----GlyPr 576  
Qy 1576 CCTCGGGTGTGAGAAATTTAGCGCAAGGACATTTTAAAGTGGATTGTTTAAAGAAC 1635  
Db 576 oAlaGly-----PrometGlyProPheAsnProGlyProPheAsnGlnG1 591  
Qy 1636 T-----CTCCAGGCCCGCCACCAAGAGGCTG-----ATCACACC 1668  
Db 591 yProProGlyGlyProProProHisGlnTyrProProGlnGlyTrpGlyAsnThrTyrPr 611  
Qy 1669 TCAGTGGGAA 1678  
Db 611 oGlnTrpGln 614  
  
RESULT 6  
FUB2 HUMAN STANDARD; PRT; 707 AA.  
ID Q92435, O00301; Q9UNT5; Q9UQH5;  
AC DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DE Far upstream element binding protein 2 (FUSE binding protein 2) (KH  
DE type splicing regulatory protein) (KSRP) (p75).  
GN FUBP2 OR KHSRP.  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., SEQUENCE OF 72-85; 123-128; 267-281; 283-291;  
RP 348-359; 472-486; 488-492; 620-625; 627-644 AND 646-651, AND FUNCTION.  
RC TISSUE=Neuroblastoma, and Retinoblastoma;  
RX MEDLINE=97282621; PubMed=9136930;  
RA Min H., Turck C.W., Nikolic J.M., Black D.L.;  
RT "A new regulatory protein, KSRP, mediates exon inclusion through an  
RT intronic splicing enhancer.";  
RL Genes Dev. 11:1023-1036(1997).  
RN [2]  
RP SEQUENCE OF 1-115 AND 570-707 FROM N.A.  
RP MEDLINE=99189245; PubMed=10087204;  
RA Ring H.Z., Vaneghi-Meyers V., Nikolic J.M., Min H., Black D.L.,  
RA Francke U.;  
RT "Mapping of the KHSRP gene to a region of conserved synteny on human  
RT chromosome 19p13.3 and mouse chromosome 17.";  
RL Genomics 56:350-352(1999).  
RN [3]  
RP SEQUENCE OF 60-707 FROM N.A., AND FUNCTION.  
RC TISSUE=B-cell lymphoma, and Skeletal muscle;  
RX MEDLINE=97094955; PubMed=8940189;  
RA Davis-Smyth T., Duncan R.C., Zheng T., Michelotti G., Levens D.;  
RT "The far upstream element-binding proteins comprise an ancient family  
RT of single-strand DNA-binding transactivators.";  
RL J. Biol. Chem. 271:31679-31687(1996).  
RN [4]











FT DOMAIN 425 492 KH 4.  
 FT DOMAIN 572 685 4 X 12 AA IMPERFECT REPEATS.  
 FT REPEAT 572 583 1.  
 FT REPEAT 618 629 2.  
 FT REPEAT 644 655 3.  
 FT REPEAT 674 685 4.  
 FT DOMAIN 7 68 GLY/PRO-RICH.  
 FT DOMAIN 69 498 GLY-RICH.  
 FT DOMAIN 499 613 ALA/GLY/PRO-RICH.  
 SQ SEQUENCE 721 AA; 74226 MW; 482C7A765C60E24A CRC64;

Alignment Scores:  
 Pred. No.: 2.45e-07 Length: 721  
 Score: 228.50 Matches: 122  
 Percent Similarity: 31.24% Conservative: 67  
 Best Local Similarity: 20.17% Mismatches: 184  
 Query Match: 7.35% Indels: 232  
 DB: 1 Gaps: 22

US-09-270-437D-5 (1-1708) x FUB2\_BAT (1-721)

QY 128 AATGGCGCGAGGGGCTTTGGCTCTCGGGGTACGCCGCCAG----- 172  
 Db 98 AsnAsnThrProAspPheGlyPheGlyGlnLysArgGlnLeuGluAspGlyAsp 117  
 QY 173 -----GGCTCACCTGTGGCAGCG 190  
 Db 118 GlnProAspSerLysLysLeuAlaSerGlnGlyAspSerIleGlySerGlnLeuGlyPro 137  
 QY 191 GGGGCCCCAGCAGCAGCAGTGGACATCCCTCGGTCTCGGTCTCGGCCAGCCAG 250  
 Db 138 IleHisProProArgThrSerMetThrGluGluTyArg-----ValProAspGly 155  
 QY 251 TATGTGGTGCCATTATGGCAGAGGAGGGGGCCACCATCCGCCAATCATCAAAACAGACC 310  
 Db 156 MetValGlyLeuIleIleGlyArgGlyGlyGluGlnIleAsnLysIleGlnGlnAspSer 175  
 QY 311 CAGTCCAGATAGACGTAGAGAGGAGAGCAGGTCCAGTCCAGTCAAAAGCCATCAGT 370  
 Db 176 GlyCysLysValGlnIle---SerProAspSerGlyGlyLeuProGluArgSerValSer 194  
 QY 371 GTGCATCCACCCCTGAGGGCTGCTCTCCGCTGTGAAGATGATCTTGGAGATTATG--- 427  
 Db 195 LeuThrGlyAlaProGluSerValGlnLysAlaLysMetMetLeuAspIleValSer 214  
 QY 428 -----CATTAAGAGGCTAAGGACACCAACCGGCT 457  
 Db 215 ArgGlyArgGlyGlyProGlyGlnPheHisAspAsnAlaAsnGlyGlyGlnAsnGly 234  
 QY 458 GACGAGGTTCCCTGAAGATCCTCGCCCATATAACTTTGTAGGGCTCTCTATTGGCAAG 517  
 Db 235 ThrValGlnGluIleMetIleProAlaGlyLys-----AlaGlyLeuValIleGlyLys 252  
 QY 518 GAAGGCGGAACCTGAAGAAGGTAGACAGATACCGAGACAAATATCACCATC----- 571  
 Db 253 GlyGlyGluThrIleLysGlnLeuGlnGluArgAlaGlyValLysMetIleLeuIleGln 272  
 QY 572 TCCTCGTTCGAGACCTTACCCTTTACACCCCTGAGAGGACCATCACTGTGAAGGGGCC 631  
 Db 273 AspGlySerGlnAsnThrAsnValAspLysProLeuArgIleIleGlyAspProTyLys 292  
 QY 632 ATCGAGAAATTTTCAGGGCGCAGCAGGAATAATGAAGAAGTTCGGGAGGCCCTATGAG 691  
 Db 293 ValGlnGlnAlaCysGluMetValMetAspIleLeuArg-----Glu 306  
 QY 692 AATGATGGCTGCATCAGCTCTCACCTGATCCCTGGCTGAACCTGGCTGTGTAGGT 751  
 Db 307 ArgAspGlnGlyGlyPheGlyAspArgAsnGluTyArgIle----- 319  
 QY 752 CTTTTCCAGCTTCATCCAGCGCAGTCCCGCGCTCCAGCAGCGTACTGGGGCTGCT 811  
 Db 320 -----SerArgValGlyGlyGly--- 325

RESULT 9

QY 812 CCTATAGCTCCTTTATGCAGGCTCCCGAGCAGGAGATGTGCAGGTATTATCCCGCC 871  
 Db 326 -----IleAspValProValProArg 332  
 QY 872 CAGCAGTGGCGCCATCATCGCAAGAAGGGCAGACATCAACACGCTCTCCGGTTT 931  
 Db 333 HisSerValGlyValValIleGlyArgSerGlyGluMetIleLysLysIleGlnAsnAsp 352  
 QY 932 GCCAGCGCTCCATC-----AAATTTGCA--- 955  
 Db 353 AlaGlyValArgIleGlnPheLysGlnAspAspGlyThrGlyProGluLysIleAlaHis 372  
 QY 956 -----CCACCGAAACACCTGACTCCAAAGTTCTGATGTT----- 991  
 Db 373 IleMetGlyProProAspArgCysGluHisAlaAlaArgIleIleAsnAspLeuGln 392  
 QY 992 ---ATCATCATCGACCGCCAGAGCCCAATT----- 1020  
 Db 393 SerLeuArgSerGlyProPro--GlyProGlyAlaProGlyMetProGlyGlyAla 412  
 QY 1021 -----CAAGGTCAGGGAAGATCTATGGCAACTCAAGGAGGAGAACTTTTGGTCCC 1075  
 Db 412 zGlyArgGlyArgGlyGlnGlyAsnTrpGlyProGlyGlyGluMetThrPhe---- 430  
 QY 1076 AAGGAGGAAGTGAAGCTGGAGACCACATACGTGTGCCAGCATCAGCAGTGGCGGTC 1135  
 Db 431 -----Ser-IleProThrHisLysCysGlyLeuVal 440  
 QY 1136 ATTGCAAAAGTGGAAAAACGGTGAACGAGTTTGCAATTTGACGGCAGCTGAGTGTA 1195  
 Db 441 IleGlyArgGlyGlyGluAsnValLysAlaIleAsnGlnGlnThrGlyAlaPheValGlu 460  
 QY 1196 GTACCAAGAGACACGACCCCT-----GATGAGAACACCGAGTTCATCGTG----- 1240  
 Db 461 IleSerArgGlnLeuProProAsnGlyAspProAsnPhenLysLeuPheValIleArgGly 480  
 QY 1241 -----AAAAATCATCGACATTTCTATCCAGTCAGTGGCTCAACGGAAGATCCGAGAC 1294  
 Db 481 SerProGlnGlnIleAspHis-----AlaLysGlnLeuIleGluGluLysIleGlu--- 497  
 QY 1295 ATCTGGCCCGAGTTTAAAGCAGCATCAGAGGACAGATTAACCGCCAGCCAGCAGCG 1354  
 Db 497 -----GlyProLeuCysPro----- 502  
 QY 1355 AGGAAGTGACACAGCCCTCCCTGTCCCTTNGAGTCCAGGACAAACACGGCAGAAATCGA 1414  
 Db 498 -----GlyProLeuCysPro----- 502  
 QY 1415 GAGTGTGTCTCCCGCAGGCTGAGATGAGTGGGAATCCGGGACACNTGGCGCGGC 1474  
 Db 503 -----ValGlyProGly 506  
 QY 1475 TGTAGATCAGGTTTCCCCACTTTGATTGAGAAAGATGTTCCAGTGAGGAACCTGATCINT 1534  
 Db 506 ----- 506  
 QY 1535 CAGCCCCAAACACCCCAATTTGCCCAACACTGTNTGCCCTCGGGGTGTCAGAAATT 1594  
 Db 507 -----ProGlyGlyProGlyProAla 513  
 QY 1595 NTAGCGCAAGGCACCTTTAAAGCTGGATTGTTAAAGAGCT----- 1636  
 Db 514 GlyProMetGlyProPheHisProGlyProPheAsnGlnGlyProGlyAlaProPro 533  
 QY 1637 -----CTCCAGGCCCCCAACAGAGGGTGG-----ATCACACT 1669  
 Db 534 HisAlaGlyGlyProProHisGlnTyProProGlnGlyTyProGlyAsnThrTyPro 553  
 QY 1670 CAGTGGGAA 1678  
 Db 554 GluTrpGln 556





Db 31 GlyGluAspGlyGlnTyrPheLeuLysValLeuIleProSerTyrAlaAlaGlySerIle 50  
Qy 266 ATTGCGAAGGAGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCAGATAGAC 325  
Db 51 IledGlyysGlyGlnThrIleValGlnLeuGlnLysGlnThrGlyAlaThrIleLys 70  
Qy 326 GTGCATAGGAGGAGAAC-----GCAGGTGCAGCTGAAAGAACCCATCACTGTGCACCTC 379  
Db 71 LeuSerLysSerLysAspPheTyrProGlyThrThrGluArgValCysLeuIleGlnGly 90  
Qy 380 ACCCTCAGGCGCTCTCTCCGCT-----TGTAAGATGATCTTGGAGATATG 427  
Db 91 ThrIleGluAlaLeuAsnAlaValHisGlyPheIleAlaGlnLysIleArgGluMetPro 110  
Qy 428 CATTAAGAGGCTAAG----- 442  
Db 111 GlnAsnValAlaLysThrGluProValSerIleLeuGlnProGlnThrThrValAsnPro 130  
Qy 443 -----GACACAAAACGGCTGACGAGGT 466  
Db 131 AspArgIleLysGlnThrLeuProSerSerProThrThrLysSerSerProSerAsp 150  
Qy 467 CCCCTG-----AAGATCTGGCCCATATAACTTTGTAGG 502  
Db 151 ProMetThrThrSerArgAlaAsnGlnValLysIleValProAsnSerThrAlaGly 170  
Qy 503 CGTCTCATGCGAAGGAGGAGAACCTGAAGAGGTAGAGCAAGATACCGAGACAAA 562  
Db 171 LeuIleIleGlyLysGlyGlyAlaThrValLysAlaIleMetGlnSerGlyAlaTtp 190  
Qy 563 ATCACCATCTCCCTGTTGCAAGCTTACCTTTACACCCCTGAGAGGACCATCATCTG 622  
Db 191 ValGlnLeuSerGlnLysProAsp-----GlyIleAsnLeuGlnGluArgValThrVal 209  
Qy 623 AAGGGGCCATCGAGATTTTCAGGGCCGAGCAGGAAATATGAAGAAAGTTTCGGAG 682  
Db 210 SerGlyGluProGluGlnAsnAlaGlyAlaValGluLeuIleGlnLysIleGlnGlu 229  
Qy 683 -----GCTATGAGAT----- 694  
Db 230 AspProGlnSerGlySerCysLeuAsnIleSerTyrAlaAsnValThrGlyProValAla 249  
Qy 695 -----GATGTGGCTGCC 706  
Db 250 AsnSerAsnProThrGlySerProTyrAlaAsnThrAlaGluValLeuProThrAlaAla 269  
Qy 707 ATGAGCTCTCACTGATCCCTGCGCTGAACTGGCTGCTGAGTGTTCCTTCCCA----- 760  
Db 270 AlaAlaAlaGlyLeuLeuGlyHisAlaAsnLeuAlaGlyValAlaAlaPheProAlaVal 289  
Qy 760 ----- 760  
Db 290 LeuSerGlyPheThrGlyAsnAspLeuValAlaIleThrSerAlaLeuAsnThrLeuAla 309  
Qy 761 -----GCTTCATCCAGC 772  
Db 310 SerTyrGlyTyrAsnLeuAsnThrLeuGlyLeuSerGlnAlaAlaIleThrGly 329  
Qy 773 GCAGTCCCGCGCTCCAGCAGCGTTACTGGGCTGCTCCCTATAGTCTCTTTATGAG 832  
Db 330 AlaLeuAlaAlaAlaAlaSerAlaAsnProAlaAlaAlaAlaAlaAsnLeuAla 349  
Qy 833 GCTCCCGAGCAGGATGTTGCGAGTGTATTATCCCGCCAGCGAGTGCGGCCCATCATC 892  
Db 350 ThrTyrAlaSerGlu-----AlaSerAlaSerGlySerThrAla 362  
Qy 893 GGCAGAAGGGGAGCAGCATCAACAGCTCTCCCGTTTCCAGCGCCCTCCATCAAGATT 952  
Db 363 GlyGlyThrAla-----GlyThrPheAlaLeuGlySerLeuAlaAla 376  
Qy 953 GCACCCCGGAAACACCTGACTCCAAAGTTCGTATGTTATCATCATCGAGCCGCGAG 1012  
Db 377 Ala-----ThrAlaAlaThrAsn 382

Qy 1013 GCCCAATTCAAGGCTCAGGAGGAATCTATGCGCAACTC-----AAGGAGGAGAAC 1063  
Db 383 GlyTyrPheGlyAlaAlaSerProLeuAlaAlaSerAlaIleLeuGlyThrGluLysSer 402  
Qy 1064 TTCCTTGGTCCCAAGGAGGAAGTGAAGTGGAGACCCACATACATGTCGCCAGCATCAGCA 1123  
Db 403 ThrAspGlySerLysAspValValGlu-----IleAlaValProGluAsnLeu 418  
Qy 1124 GCTGCCCGGCTCATGCGAAGGTGGAACACGTTGCAACGTTGCGAATTTGACGGCA 1183  
Db 419 ValGlyAlaIleLeuGlyLysGlyLysThrLeuValGluTyrGlnGluLeuThrGly 438  
Qy 1184 GCTGAGGTGGTAGTACCAAGA-----GACCAGACCCCTGATGAGAACGACCATCATC 1237  
Db 439 AlaArgIleGlnIleSerLysLysGlyGluPheValProGlyThrArgAsnArgLys--- 457  
Qy 1238 GTGAATATCATCGACATTTCTATGCCAGTCAGATGGCTCAA 1279  
Db 458 ValThrIleThrGlyThrProAlaAlaThrGlnAlaAlaGln 471  
RESULT 11  
NOAL HUMAN  
ID NOAL HUMAN STANDARD; PRT; 510 AA.  
AC P51513;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE RNA-binding protein Nova-1 (Neuro-oncological ventral antigen 1)  
DE (Onconeural ventral antigen-1) (Paraneoplastic R1 antigen) (Ventral  
DE neuron-specific protein 1).  
GN NOAL.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).  
RC TISSUE=Cerebellum, and Hippocampus;  
RX MEDLINE=9400830; PubMed=8398153;  
RA Buckanovich R.J., Posner J.B., Darnell R.B.;  
RT "Nova, the paraneoplastic R1 antigen, is homologous to an RNA-binding  
RT protein and is specifically expressed in the developing motor  
RT system".  
RL Neuron 11:657-672(1993).  
RN [2]  
RP SEQUENCE OF 1-34 FROM N.A.  
RC TISSUE=Fetal brain;  
RA Dmitrenko V.V., Garifulin O.M., Shostak K.A., Smikodub A.I.,  
RL Kavsan V.M.;  
RN Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 423-495.  
RX MEDLINE=99148126; PubMed=10368286;  
RA Lewis H.A., Chen H., Edo C., Buckanovich R.J., Yang Y.Y.-L.,  
RA Musunuru K., Zhong R., Darnell R.B., Burley S.K.;  
RT "Crystal structures of Nova-1 and Nova-2 K-homology RNA-binding  
RT domains".  
RL Structure 7:191-203(1999).  
CC -!- FUNCTION: May regulate RNA splicing or metabolism in a specific  
CC subset of developing neurons.  
CC -!- SUBCELLULAR LOCATION: Nuclear.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=3;  
CC Name=1;  
CC IsoId=P51513-1; Sequence=Displayed;  
CC Name=2; Synonyms=tumor;  
CC IsoId=P51513-2; Sequence=VSP\_002841, VSP\_002843, VSP\_002844;  
CC Name=3;  
CC IsoId=P51513-3; Sequence=VSP\_002842;  
CC -!- TISSUE SPECIFICITY: Brain.  
CC -!- DISEASE: Autoantigen in the paraneoplastic opsoclonus myoclonus  
CC ataxia (POMA), a paraneoplastic neurological syndrome/disorder









PCB4\_HUMAN STANDARD; PRT; 403 AA.  
P57723;  
16-OCT-2001 (Rel. 40, Created)  
16-OCT-2001 (Rel. 40, Last sequence update)  
10-OCT-2003 (Rel. 42, Last annotation update)  
Poly(rC)-binding protein 4 (Alpha-CP4).  
PCBP4  
Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP MEDLINE=20396135; PubMed=10936052;  
RA Makeyev A.V.; Liebhauer S.A.;  
RT "Identification of two novel mammalian genes establishes a subfamily  
of KH-domain RNA-binding proteins.";  
RL Genomics 67:301-316(2000).  
[2]  
SEQUENCE FROM N.A.  
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H., Sugawara M.,  
RA Wgatsuma M., Hosori T., Kaku Y., Kodaira H., Kondo H., Takiguchi S.,  
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,  
RA Nanomiya K., Iwayanagi T.;  
RT "NEDO human cDNA sequencing project.";  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
[3]  
SEQUENCE FROM N.A.  
RP TISSUE=Lung;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnerch A., Schein J.B., Jones S.J.W., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
CC -!- FUNCTION: Single-stranded nucleic acid binding protein that binds  
preferentially to oligo dC (By similarity).  
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).  
CC -!- SIMILARITY: Contains 3 KH domains.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; AF176330; AAG09241.1; -  
DR EMBL; AK023993; BAB14761.1; -  
DR EMBL; BC003008; AAH03008.1; -  
DR EMBL; BC004153; AAH04153.1; -  
DR Genew; HGNC:8652; PCBP4.  
DR GO; GO:0005634; C:nucleus; ISS.  
DR GO; GO:0003677; F:DNA binding; ISS.

DR GO; GO:0003723; F:RNA binding; NAS.  
DR GO; GO:0016071; P:mRNA metabolism; ISS.  
DR InterPro; IPR004087; KH\_dom.  
DR InterPro; IPR004088; KH\_type\_1.  
DR Pfam; PF00013; KH; 3.  
DR SMART; SMO0322; KH; 3.  
DR PROSITE; PS00084; KH\_TYPE\_1; 3.  
KW Nuclear protein; RNA-binding; Ribonucleoprotein; DNA-binding;  
KW Repeat.  
FT DOMAIN 17 67 KH 1.  
FT DOMAIN 101 154 KH 2.  
FT DOMAIN 241 293 KH 3.  
SQ SEQUENCE 403 AA; 41481 MW; 3D99F762A9471265 CRC64;  
  
Alignment Scores:  
Pred. No.: 1,11e-06 Length: 403  
Score: 217.50 Matches: 80  
Percent Similarity: 42.63% Conservatives: 53  
Best Local Similarity: 25.64% Mismatches: 114  
Query Match: 6.99% Indels: 65  
DB: 1 Gaps: 11  
  
US-09-270-437D-5 (1-1708) x PCB4\_HUMAN (1-403)  
QY 215 GTGACATCCCTTCGGCTCTGCTGCCACCCAGTATGTGGTGCCATTATTGGCAAG 274  
Db 16 lIeThrLeuThrLeuArgMetLeuMethHisGlyLysGluValGlySerIleIleGlyLys 35  
QY 275 GAGGGGGCCCATCCGCAACATCACAACAGACCCAGTCCAGATAGACGTGCATAGG 334  
Db 36 LysGlyGluThrValLysArgIleArgGluInSerSerAlaArgIle----- 51  
QY 335 AAGGAGAACGGAGTGCAGCTGAAAGCCATCATGTGTGCATCCACCCCTGAGGGTGC 394  
Db 52 -----ThrIleSerGluGlySerCysProGluArgIle 62  
QY 395 TCCTCCGCTTGTAAATGATCATCTTGAGATTATGCTAAGAGGCTAAGACACCAACG 454  
Db 63 ThrThrIleThrGlySerThrAlaAlaValPheHisAlaValSerMetIleAlaPheLys 82  
QY 455 GCTGACGAG-----GlyAlaAlaProAlaAlaGlyGlyAsnValSerArgProVal 102  
QY 467 CCCCTGAAGATCCTGCCCATTAATTAATTTGTAGGCGCTCTCATTTGGCAGGAAGACGG 526  
Db 103 ThrLeuArgLeuValIleProAlaSerGlnCysGlySerLeuIleGlyLysAlaGlyThr 122  
QY 527 AACCTGAAGAGGTAGACGAGATACCGAGACAAAAATCACCATCTCCTCGTTGCAAGAC 586  
Db 123 LysIleLysGluIleArgGluThrThrGlyAlaGlnValGlnValAlaGly----- 140  
QY 587 CTTACCCCTTTACACCCCTGAGAGGACCATCTGTGAAGGGGGCCATCGAGAATTGTGTC 646  
Db 141 LeuLeuProAsnSerThrGluArgAlaValThrValSerGlyValProAspAlaIle 160  
QY 647 AGGGCCGACGAGAAATAATGAAGAACTTCGGAGGCGCTATGAGATGATGCTGCTGCC 706  
Db 161 LeuCysValArgGlnIleCysAlaValIleLeuGluSerProLysGly---AlaThr 179  
QY 707 ATGAGCTCTCACCTGATCCCTGCCCTGACCTGGCTGCTGTA----- 748  
Db 180 IleProTyHis-----ProSerLeuSerLeuGlyThrValLeuLeuSerAlaAsnGln 197  
QY 749 -----GGTCTTTTCCAGCT----- 763  
Db 198 GlyPheSerValGlnGlnGlyThrGlyAlaValThrProAlaGluValThrLysLeuGln 217  
QY 764 -----TCATCCAGCGCAGTCCCG---CCGCTCCAGCAGCGTGTACTGGGCTGCTCC 814  
Db 218 GlnLeuSerSerHisAlaValProPheAlaThrProSerValValProGlyLeuAspPro 237  
QY 815 TATAGCTCCTTTATGCAGGCTCCCGAGCAGGAGATGGTGCAGGTGTTTATCCCGCCGAG 874

Db 238 GlyThr-----GlnThrSerSerGlnGluPheLeu-----ValProAsnAsp 251  
 QY 875 GCAGTGGCGGCATCATCGGCAAGAGGGGAGACATCAACAGCTCTCCGGTTTGCC 934  
 Db 252 LeuileGlyCysValileGlyArgGlnGlySerLysileSerGluileA-glnMetSer 271  
 QY 935 AGCGCTCCATCAAGATTGACCCAGCCGCAACACCTGACTCCAAAGTTCTATGGTTATC 994  
 Db 272 GlyAlaHisIleIleGlyGly--AsnGlnAlaGluGlyAlaGlyGluArgHisValThr 290  
 QY 995 ATCACTGACCGCCAGAGGCCCAATTCAGGCTCAG 1030  
 Db 291 IleThrGlySerProValSerIleAlaLeuAlaGln 302

## RESULT 14

PCB4\_MOUSE  
 ID PCB4\_MOUSE STANDARD; PRT; 403 AA.  
 AC F57724;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Poly(rC)-binding protein 4 (Alpha-CP4).  
 GN PCBp4  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20396135; PubMed=10936052;  
 RA Makeyev A.V., Liebraber S.A.;  
 RT "Identification of two novel mammalian genes establishes a subfamily  
 of KH-domain RNA-binding proteins.";  
 RL Genomics 67:301-316(2000).  
 CC -!- FUNCTION: Single-stranded nucleic acid binding protein that binds  
 preferentially to oligo dC (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).  
 CC -!- TISSUE SPECIFICITY: Ubiquitous.  
 CC -!- SIMILARITY: Contains 3 KH domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF176328; AAC09239.1; .  
 CC MGD; MGI:1890471; Pcbp4.  
 CC InterPro; IPR004087; KH dom.  
 CC InterPro; IPR004088; KH\_type\_1.  
 CC Pfam; PF00013; KH; 3.  
 CC SMART; SM00322; KH; 3.  
 CC PROSITE; PS50084; KH TYPE 1; 3.  
 CC Nucleic protein; RNA-Binding; Ribonucleoprotein; DNA-binding;  
 KW Repeat.  
 FT DOMAIN 17 67 KH 1.  
 FT DOMAIN 101 154 KH 2.  
 FT DOMAIN 241 293 KH 3.  
 SQ SEQUENCE 403 AA; 41398 MW; D4DBFF490CFE90A CRC64;

## Alignment Scores:

Pred. No.: 1-73e-06 Length: 403  
 Score: 214.50 Matches: 79  
 Percent Similarity: 42.31% Conservative: 53  
 Best Local Similarity: 25.32% Mismatches: 115  
 Query Match: 6.90% Indels: 65  
 DB: 1 Gaps: 11

US-09-270-437D-5 (1-1708) x PCB4\_MOUSE (1-403)

QY 215 GTGGACATCCCTTCGGCTCCTGGTCCGCCACCCAGTATGTGGTGCCATTATTGGCAAG 274  
 Db 16 IleThrLeuThrLeuArgMetLeuMethIsglyysGluValGlySerIleileGlyys 35  
 QY 275 GAGGGGGCCACCATCCGCAACATCACAACAGAGCCAGTCCAGATAGACGTGCATAGG 334  
 Db 36 LysGlyGluThrValIleArgIleArgGluGlnSerSerAlaArgile----- 51  
 QY 335 AAGGAGACGAGGTGCAGCTGAAAACCCATCAGTGTGCACTCCACCCGTAGGGGTGC 394  
 Db 52 -----ThrIleSerGluGlySerCysProGluArgile 62  
 QY 395 TCCTCCGCTTGTAGATGATCTTGGAGATTATGATAAGAGGCTAAGGACACCAAAACG 454  
 Db 63 ThrIleThrGlySerThrAlaAlaValPheHisAlaValSerMetIleAlaPheIys 82  
 QY 455 GCTGACGAG-----GTT 466  
 Db 83 LeuAspGluAspLeuCysAlaAlaProAlaAsnGlyGlySerValSerArgProProVal 102  
 QY 467 CCCCTGAAGATCCCTGGCCCAATAAATCTTGTAGGGCTCTCATTTGGCAAGAGGACGG 526  
 Db 103 ThrLeuArgLeuValIleProAlaSerGlnCysGlySerLeuIleGlyLysAlaGlyThr 122  
 QY 527 AACCTGAAGAAGTAGAGCAAGATACCGAGACAAAATCACCATCTCTCTCGTTCAAGAC 586  
 Db 123 LysIleLysGluIleArgGluThrThrGlyAlaGlnValGlnValAlaGly-----Asp 140  
 QY 587 CTTACCCCTTTACACCCCTGAGGAGGACCATCATCTGTAAGGGGGCCATCGAGAATTGTTC 646  
 Db 141 LeuLeuProAsnSerThrGluArgAlaValThrValSerGlyValProAspAlaIle 160  
 QY 647 AGGCGCAGAGGAATAATCAAGAAAGTTCCGAGGCTATGAGATGATGTCGCTGCC 706  
 Db 161 LeuCysValArgGlnIleCysAlaValIleLeuGluSerProProLysGly---AlaThr 179  
 QY 707 ATGAGCTCTCACTGATCCCTGGCTGCACTGCTGCTGTA----- 748  
 Db 180 IleProTyHis-----ProSerLeuSerLeuGlyThrValLeuLeuSerAlaAsnGln 197  
 QY 749 -----GGTCTTTTCCAGCT----- 763  
 Db 198 GlyPheSerValGlnGlyGlnTyGlyAlaValThrProAlaGluValThrLysLeuGln 217  
 QY 764 -----TCATCCAGCAGGTCGCG-----CCGCTCCAGAGCGCTTACTGGGTGCTCCC 814  
 Db 218 GlnLeuSerGlyHisAlaValProPheAlaSerProSerValValProGlyMetAspPro 237  
 QY 815 TATAGCTCTTTATGACAGGCTCCCGAGCAGGATGTCGAGGTGTTATCCCGCCAG 874  
 Db 238 SerThr-----GlnThrSerSerGlnGluPheLeu-----ValProAsnAsp 251  
 QY 875 GCAGTGGCGGCATCATCGGCAAGAGGGGAGACATCAACAGCTCTCCCGTTTGCC 934  
 Db 252 LeuileGlyCysValileGlyArgGlnGlySerLysileSerGluileA-glnMetSer 271  
 QY 935 AGCGCTCCATCAAGATTGACCCAGCCGCAACACCTGACTCCAAAGTTCTATGGTTATC 994  
 Db 272 GlyAlaHisIleIleGlyGly--AsnGlnAlaGluGlyAlaGlyGluArgHisValThr 290  
 QY 995 ATCACTGACCGCCAGAGGCCCAATTCAGGCTCAG 1030  
 Db 291 IleThrGlySerProValSerIleAlaLeuAlaGln 302

## RESULT 15

Y475\_ARATH  
 ID Y475\_ARATH STANDARD; PRT; 606 AA.  
 AC P58223; Q49507; Q8H0Y7;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Putative nucleic acid binding protein At4g18375.  
 GN AT4G18375 OR F28U12.2.



```

Db      118  AspaLalleValAlaSerAspGluGluAsnAsnThrLysThrAsnValAspArgAsp 137
QY      461  GAGGTTCCTCCCTGAGATCCTCGCCCATATAACTTTGTAGGGCTCNCATTGGCAGAGAA 520
Db      138  AsnLysGluCysArgLeuValProPheSerGlnSerSerLeuIleGlyLysAla 157
QY      521  GGACGGAACCTGAAGAAGGTAGACAAGATACCGAGACAAAAATCACCATCTCTCGTTG 580
Db      158  GlyGluAsnIleLysArgIleArgArgThrArgAlaSerValLysValValSer-- 176
QY      581  CAAGACCTTACCCTTTACACCTT-----GAGAGACCACATC 616
Db      177  LysAspValSer-----AspProSerHisValCysAlaMetGluThrAspAsnValVal 194
QY      617  ACTGTGAAGGG-----GCCATCGAGAATTTGTCAGGGCCGAGCAGGAAATAATG 667
Db      195  ValIleSerGlyGluProGluSerValLysGlnAlaLeuPheAlaValSerAlaIleMet 214
QY      668  AAGAAGTTCCG-----GAGCCTATGAGAAATGATGTGGTGCC 706
Db      215  TyrLysIleAsnProArgGluAsnIleProLeuAspSerThrSerGlnAspValProAla 234
QY      707  ATGAGC-----712
Db      235  AlaSerValIleValProSerAspLeuSerAsnSerValTyrProGlnThrGlyPheTyr 254
QY      713  -----TCTACCTG-----ATCCTGCG-----CTGAACCTGGGT 742
Db      255  SerAsnGlnAspHisIleLeuGlnGlnGlyAlaGlyValProSerTyrPheAsnAlaLeu 274
QY      743  GCTGTAGTCTTTTCCCA-----GCTTCATCCAGCGCAGTCCCGCGCTCCACGACG 796
Db      275  SerValSerAspPheGlnGlyTyrAlaGluThrAlaAlaAsnProValProValPheAla 294
QY      797  GTTACTGGGGTGCTCCTATAGTCTCCTTTATCAGGCTCCCGAGCAGGAGATGGTG--- 853
Db      295  SerSerLeuProValThrHisGlyPheGlyGlySerSerArgSerGluGluLeuValPhe 314
QY      854  CAGGTGTTATCCCGCCAGGAGTGGGGCCCATCATCGGCAAGAGGGGAGCAGCATC 913
Db      315  LysValLeuCysProLeuCysAsnIleMetArgValIleGlyLysGlySerThrIle 334
QY      914  AAACAGCTCTCCGGTTTGCAGCGCTCCATCAGATTGCACCCCGCAACACCTGAC 973
Db      335  LysArgIleArgGluAlaSerGlySerCysIle-----GluValAsnAsp 349
QY      974  TCCAAAGTTCTG-----ATGGTTATCATCTGAGCCGCGCAGAGGCC 1015
Db      350  SerArgThrLysCysGlyAspAspGluCysValIleValThrAlaThrGluSerPro 369
QY      1016  CAATTCAAGGCTCAGGAGAAATCTATGGCAAACTCAAGGAGGAACTTCTTTGGTCCC 1075
Db      370  AspAspMetLysSerMetAlaValGluAlaValLeuLeuGlnGluThrIleAsnAsp 389
QY      1076  AAGGAGAGTGAAGTGGAGACCCACATACGTGTGCCAGCATCAGCTGCGCGGTC 1135
Db      390  GluAspAlaGluAsnValLysMetGlnLeuLeuValSerSerLysValIleGlyCysVal 409
QY      1136  ATTGGCAAGGTGAAACCGGTGACGAGTTGCAGAAATTTGACGCGCAGCTGAGTGGTA 1195
Db      410  IleGlySerGlySerValIleAsnGluIleArgLysArgThrAsnAlaAsnIleCys 429
QY      1196  GTACCAAGAGACCAACCCCTGATGAGAACGACAGGTCTCGTGAATAATCATCGACAT 1255
Db      430  IleSerLysGlyLysLysAspAsp-----LeuValGluValSerGlyGlu 444
QY      1256  TTCATGCCAGTCAGATGGCTCAACGGAGATCCGAGACATCCTGGCCCGAGTT----- 1309
Db      445  Val-----SerSerValArgAspAlaLeuIleGlnIleValLeu 457
QY      1310  ---AAGCAGCAGCATCAGAAGGACAGAGTAACAGGCCCGCCAGGACGAGGAAGTGACCA 1366

```

```

Db      458  ArgLeuArgGluAspValLeuGlyAspLysAspSerValAlaThrArgLys----- 474
QY      1367  GCCCCTCCCTGTCCTCTTGAG-----TCCAGGACAAACACGGGACAGAAATCGAGAGTGT 1420
Db      475  ---ProProAlaArgThrAspAsnCysSerPheLeuSerGlySerSerAsnAlaGlyTyr 493
QY      1421  GCTCTCCCC-----GGC 1432
Db      494  ThrLeuProSerPheMetSerSerMetAlaSerThrSerGlyPheHisGlyTyrGlySer 513
QY      1433  AGGCCTGAGATGACTGGGAATCCGGGACACNTGGCGCGGCTGTAGATCAGGTTTGGCC 1492
Db      514  PheProAlaGlyAspAsnValLeuGlySerThrGlyProTyrSerTyrGlyArgLeuPro 533
QY      1493  ACT 1495
Db      534  Ser 534

```

Search completed: July 13, 2004, 12:09:27  
Job time : 42.5 secs